ANNALS OF MATHEMATICS
Vol. 38, No. 3, July, 1937

REPRESENTATIONS INDUCED IN AN INVARIANT SUBGROUP

BY A. H. CLIFFORD

(Received January 19, 1937; revised April 30, 1937)

The object of this paper is to investigate the representation $\mathfrak{A}_H$ induced by an irreducible representation $\mathfrak{A}_G$ of a group G in an invariant subgroup H of G. In §1 it is shown that $\mathfrak{A}_H$ is either itself irreducible or is fully reducible into conjugate irreducible representations of H. In §2 it is shown that $\mathfrak{A}_G$ is imprimitive unless all the irreducible components of $\mathfrak{A}_H$ are equivalent. In fact, if $\mathfrak{R}$ is the representation space of $\mathfrak{A}_G$, and hence also of $\mathfrak{A}_H$, and if we lump together all equivalent subspaces of $\mathfrak{R}$ under $\mathfrak{A}_H$, then the resulting subspaces $\mathfrak{R}_1, \mathfrak{R}_2, \cdots, \mathfrak{R}_m$ constitute a system of imprimitivity of $\mathfrak{A}_G$. If we define G' to be the subgroup of G leaving one of these invariant, say $\mathfrak{R}_1$, then the component of $\mathfrak{A}_{G'}$ in $\mathfrak{R}_1$ is an irreducible representation $\mathfrak{A}'_{G'}$ of G', and $\mathfrak{A}_G$ is expressible very simply in terms of $\mathfrak{A}'_{G'}$.

These results hold for any group G and any ground-field P. In §§3–5, however, we make the assumption that P is algebraically closed. In §3 it is found that $\mathfrak{A}'_{G'}$ is the direct product of two irreducible projective representations of G', one of which is actually a projective representation Γ of the factor-group G'/H. In §4 some progress is made on the question of whether or not a given irreducible representation of H can be embedded in some irreducible representation of G, and in §5 we consider all possible ways of doing this. Two irreducible representations $\mathfrak{A}_G$ and $\mathfrak{B}_G$ of G are said to be associate if $\mathfrak{A}_H$ and $\mathfrak{B}_H$ have an irreducible component in common; associates differ only in the projective representation Γ of G'/H mentioned above.

In the case when the factor-group G/H is a finite cyclic group of order k, associates can be described as differing from each other only by a factor which is a one-dimensional representation of G/H, and hence just a k^{th} root of unity. In the simplest case of all, when H is of index two in G, $\mathfrak{A}_G$ has just one associate $\mathfrak{A}^*_G$ (besides itself) differing from $\mathfrak{A}_G$ only in that we change the sign of the matrices corresponding to elements of G not in H. The situation may then be described as follows. If $\mathfrak{A}_G$ is not equivalent to $\mathfrak{A}^*_G$, then $\mathfrak{A}_H$ is irreducible. If $\mathfrak{A}_G$ is equivalent to $\mathfrak{A}^*_G$, then $\mathfrak{A}_H$ decomposes into two inequivalent (conjugate) irreducible components. If $\mathfrak{B}_G$ is another irreducible representation of G equivalent to neither $\mathfrak{A}_G$ nor $\mathfrak{A}^*_G$, then $\mathfrak{B}_H$ can have no irreducible component in common with $\mathfrak{A}_H(= \mathfrak{A}^*_H)$.

Virtually all of this theory is known in the case of a finite group G, and the greater part of it goes back to Frobenius. For the decomposition of $\mathfrak{A}_H$ into conjugates we must refer to Frobenius' original paper.[1] For the results of §2—

[1] G. Frobenius, *Über Relationen zwischen den Charakteren einer Gruppe und denen ihrer Untergruppen*, (S.-B. preuss. Akad. Wiss. Berlin 1898, 501–515), p. 506.

the imprimitivity of $\mathfrak{A}_G$ and its generation by $\mathfrak{A}'_{G'}$—we may refer to Speiser[2] and van der Waerden.[3] Theorem 2 below is a direct extension of Speiser's Theorem 172 on p. 195 to arbitrary linear groups in any field. Frobenius[4] knew the theory of associates, as described above, when G is the symmetric and H the alternating group, and Burnside[5] has a close approximation to it when H is an invariant subgroup of prime index in any finite group G. Brauer[6] was in full possession of these results for the case when G is the full orthogonal and H the proper orthogonal group, and Professor Weyl—to whom I am indebted for interesting me in the problem—for the case of any group G containing a subgroup H of index two.

Nakayama and Shoda[7] have obtained the results of §§1–3 in the case of an irreducible representation of a finite group G by semi-linear transformations, H being the invariant subgroup of G consisting of those elements of G which effect the identical automorphism in the ground-field. In the concluding section, §6, I shall indicate briefly the modifications necessary to carry over the results of §§1 and 2 to this more general case. The results are valid for any group G, any invariant subgroup H of G, and any ground-field P.

1. Full Reducibility of $\mathfrak{A}_H$

We begin with any abstract group G, any invariant subgroup H of G, and any ground-field P. The elements of G will be denoted by $r, s, t, \cdots$, those of H by $u, v, \cdots$, and those of P by small Greek letters. If $u \to A(u)$ is any representation $\mathfrak{A}_H$ of H by matrices in P, $A(u)$ indicating the matrix corresponding to the element u of H, and if r is any fixed element in G, then the representation $\mathfrak{A}_H^*$ of H defined by $u \to A^*(u) = A(r^{-1}ur)$ will be called a *conjugate* of $\mathfrak{A}_H$ relative to G. $\mathfrak{A}_H$ and $\mathfrak{A}_H^*$ evidently have the same degree (number of rows and columns of the representing matrices), and if $\mathfrak{A}_H$ is fully reducible so also is $\mathfrak{A}_H^*$, and the irreducible components of $\mathfrak{A}_H^*$ are conjugates of those of $\mathfrak{A}_H$.

If $s \to A(s)$ is any representation $\mathfrak{A}_G$ of G then we shall say that $\mathfrak{A}_G$ *induces* the representation $\mathfrak{A}_H$ of H defined by $u \to A(u)$. We propose to study the representations of H induced by *irreducible* representations of G.

THEOREM 1. *If $\mathfrak{A}_G$ is any irreducible representation of a group G in any field* P,

---

[2] A. Speiser, *Theorie der Gruppen von endlicher Ordnung*, (second edition, Berlin, 1927), pp. 191–199.

[3] B. L. van der Waerden, *Gruppen von linearen Transformationen*, (Ergebnisse der Math., v. 4, 1935), p. 31 and pp. 75–79.

[4] G. Frobenius, *Über die Charaktere der alternirenden Gruppe*, (S.-B. preuss. Akad. Wiss. Berlin 1901, 303–315), §2.

[5] W. Burnside, *Theory of Groups of Finite Order*, (second edition, Cambridge, 1911), Theorem VI on p. 337.

[6] R. Brauer, *Über die Darstellung der Drehungsgruppe durch linearer Substitutionen*, (Dissertation 1925), pp. 20–23.

[7] T. Nakayama and K. Shoda, *Über die Darstellung einer endliche Gruppe durch halblineare Transformationen*, (Jap. Jour. of Math. 12, 1936, 109–122), pp. 117–120.

and H is an invariant subgroup of G, then the representation $\mathfrak{A}_H$ of H induced by $\mathfrak{A}_G$ is either itself irreducible or is fully reducible into irreducible components all of the same degree. If $\mathfrak{A}_H^{(1)}$ is any irreducible component of $\mathfrak{A}_H$ then all the other irreducible components of $\mathfrak{A}_H$ are conjugates of $\mathfrak{A}_H^{(1)}$ relative to G, and every such conjugate of $\mathfrak{A}_H^{(1)}$ must occur in the decomposition of $\mathfrak{A}_H$.

PROOF: Let $\mathfrak{R}$ be the representation space of $\mathfrak{A}_G$ (and $\mathfrak{A}_H$). If x is a vector in $\mathfrak{R}$ we shall write sx for $A(s)x$, and if $\mathfrak{S}$ is a subspace of $\mathfrak{R}$ we shall denote by $s\mathfrak{S}$ the subspace of $\mathfrak{R}$ consisting of all vectors sx of $\mathfrak{R}$ as x ranges over $\mathfrak{S}$.

If $\mathfrak{S}$ is any subspace of $\mathfrak{R}$ invariant under $\mathfrak{A}_H$, and r is a fixed element of G, then the space $r\mathfrak{S}$ is also invariant under $\mathfrak{A}_H$. For if u is any element of H, then

$$u \cdot r\mathfrak{S} = r \cdot r^{-1}ur\mathfrak{S} \subseteq r\mathfrak{S}$$

since $r^{-1}ur$ is in H, and $u\mathfrak{S} \subseteq \mathfrak{S}$ for all u in H. Since the matrix $A(r)$ is non-singular, the spaces $\mathfrak{S}$ and $r\mathfrak{S}$ have the same dimension (and the inclusions $\subseteq$ might as well have been replaced by equalities). In fact, if $e_1, e_2, \cdots, e_n$ is a basis of $\mathfrak{S}$, then $re_1, re_2, \cdots, re_n$ is a basis of $r\mathfrak{S}$. If we set

$$ue_i = \sum_j \alpha_{ij}(u)e_j,$$

then

$$u \cdot re_i = \sum_j \alpha_{ij}(r^{-1}ur)re_j.$$

Hence, under this choice of coordinates in $r\mathfrak{S}$, u induces precisely the same linear transformation in $r\mathfrak{S}$ that $r^{-1}ur$ induces in $\mathfrak{S}$: the representation of H induced by $\mathfrak{A}_H$ in the invariant subspace $r\mathfrak{S}$ is a conjugate of that induced by $\mathfrak{A}_H$ in $\mathfrak{S}$.

If $\mathfrak{S}$ is irreducible under $\mathfrak{A}_H$ then so also is $r\mathfrak{S}$. For if $\mathfrak{T}$ were a proper invariant subspace $\neq (0)$ of $r\mathfrak{S}$ then $r^{-1}\mathfrak{T}$ would be a proper invariant subspace $\neq (0)$ of $\mathfrak{S}$.

Select now any subspace $\mathfrak{S} \neq (0)$ of $\mathfrak{R}$, invariant and irreducible under $\mathfrak{A}_H$. If $\mathfrak{S} = \mathfrak{R}$ then $\mathfrak{A}_H$ is irreducible. If $\mathfrak{S} \neq \mathfrak{R}$ there must exist an element r_2 in G such that $r_2\mathfrak{S} \neq \mathfrak{S}$; otherwise $\mathfrak{S}$ would be invariant under $\mathfrak{A}_G$, contrary to the assumption that $\mathfrak{A}_G$ is irreducible. The intersection $\mathfrak{S} \cap r_2\mathfrak{S}$ of $\mathfrak{S}$ with $r_2\mathfrak{S}$ is a proper invariant subspace of $r_2\mathfrak{S}$, and since $r_2\mathfrak{S}$ is irreducible we conclude that

$$\mathfrak{S} \cap r_2\mathfrak{S} = (0).$$

Their sum $\mathfrak{S} + r_2\mathfrak{S}$ is invariant under $\mathfrak{A}_H$, and if it is $\neq \mathfrak{R}$ there must exist an element r_3 in G such that $r_3\mathfrak{S}$ is not contained in it. It then follows that

$$(\mathfrak{S} + r_2\mathfrak{S}) \cap r_3\mathfrak{S} = (0),$$

since the left-hand member is a proper invariant subspace of the irreducible invariant subspace $r_3\mathfrak{S}$.

Continuing in this fashion, we construct a sequence of elements $r_1(= 1)$, r_2, r_3, $\cdots$ of G such that

$$(r_1\mathfrak{S} + r_2\mathfrak{S} + \cdots + r_{i-1}\mathfrak{S}) \cap r_i\mathfrak{S} = (0), \qquad (i = 2, 3, \cdots).$$

But this means that the spaces $r_i\mathfrak{S}$ are mutually independent, and since they all have the same positive dimension the total space $\mathfrak{R}$ must be exhausted after a finite number of steps:

$$\mathfrak{R} = \mathfrak{S} + r_2\mathfrak{S} + \cdots + r_h\mathfrak{S}.$$

We thus succeed in decomposing $\mathfrak{R}$ into a sum of mutually independent subspaces, invariant and irreducible under $\mathfrak{A}_H$. We have already observed that the irreducible component of $\mathfrak{A}_H$ in each $r_i\mathfrak{S}$ is a conjugate relative to G of that in $\mathfrak{S}$. That every such conjugate must occur follows from the arbitrariness in the choice of r_2; for the only restriction on r_2 was $r_2\mathfrak{S} \neq \mathfrak{S}$, while if $r_2\mathfrak{S} = \mathfrak{S}$ the two conjugates are equivalent.

As an immediate corollary of this theorem, if $\mathfrak{A}_G$ is any fully reducible representation of G, and if H is any subgroup of G that can be reached by a normal series from G, then $\mathfrak{A}_H$ is also fully reducible.

If H is of finite index k in G, say

$$G = r_1H + r_2H + \cdots + r_kH, \qquad (r_1 = 1),$$

then the number of irreducible components of $\mathfrak{A}_H$ cannot exceed k. For the spaces $r_i\mathfrak{S}$ are merely permuted among themselves by any s in G, and hence their sum must be the total space $\mathfrak{R}$. They are not a system of imprimitivity of $\mathfrak{A}_G$, however, since they will not in general be linearly independent.

2. Imprimitivity of $\mathfrak{A}_G$

A representation $\mathfrak{A}_G$ of G is said to be *imprimitive* if it is possible to decompose its representation space $\mathfrak{R}$ into a direct sum

$$\mathfrak{R} = \mathfrak{R}_1 + \mathfrak{R}_2 + \cdots + \mathfrak{R}_m$$

in such a way that the spaces $\mathfrak{R}_i$ are permuted bodily among themselves by any transformation $A(s)$ in $\mathfrak{A}_G$. The spaces $\mathfrak{R}_i$ are said to constitute a *system of imprimitivity* of $\mathfrak{A}_G$.

Continuing with the situation described in Theorem 1, let $\mathfrak{A}_H^{(1)}, \mathfrak{A}_H^{(2)}, \cdots, \mathfrak{A}_H^{(m)}$ be a complete list of the inequivalent irreducible representations of H occurring in the decomposition of $\mathfrak{A}_H$. Choose a coordinate system in the representation space $\mathfrak{R}$ of $\mathfrak{A}_G$ which exhibits the full reduction of $\mathfrak{A}_H$, and for each $i(=1, 2, \cdots, m)$ let $\mathfrak{R}_i$ be the sum of all the equivalent irreducible invariant subspaces of $\mathfrak{R}$ belonging to $\mathfrak{A}_H^{(i)}$.

THEOREM 2. *$\mathfrak{R}_1, \mathfrak{R}_2, \cdots, \mathfrak{R}_m$ constitute a system of imprimitivity of $\mathfrak{A}_G$. Each of the inequivalent irreducible components $\mathfrak{A}_H^{(i)}$ of $\mathfrak{A}_H$ occurs in $\mathfrak{A}_H$ the same number of times. If this number is l, and if the degree of each $\mathfrak{A}_H^{(i)}$ is n, then the degree of $\mathfrak{A}_G$ is lmn.*

PROOF: In accordance with the decomposition

$$\mathfrak{R} = \mathfrak{R}_1 + \mathfrak{R}_2 + \cdots + \mathfrak{R}_m$$

we write the matrices $A(s)$ of $\mathfrak{A}_G$ in the form

$$A(s) = \left\| \begin{matrix} A_{11}(s) & \cdots & A_{1m}(s) \\ \cdots & \cdots & \cdots \\ A_{m1}(s) & \cdots & A_{mm}(s) \end{matrix} \right\|.$$

For elements u in H, the matrices $A_{ij}(u)$ with $i \neq j$ are all zero, and the representation $u \to A_{ii}(u)$ of H is a multiple of the single irreducible representation $\mathfrak{A}_H^{(i)}$.

Let s be an arbitrary but fixed element of G. Since $s^{-1}us$ is in H for all u in H, the equation

$$A(u)A(s) = A(s)A(s^{-1}us)$$

leads to

$$A_{ii}(u)A_{ij}(s) = A_{ij}(s)B_{jj}(u), \qquad (i, j = 1, 2, \cdots, m),$$

where

$$B_{jj}(u) = A_{jj}(s^{-1}us).$$

The representation $u \to B_{jj}(u)$ is a conjugate of $u \to A_{jj}(u)$, and is therefore a multiple of a conjugate $\mathfrak{B}_H^{(j)}$ of $\mathfrak{A}_H^{(j)}$.

This equation holds for all u in H. By a simple extension of Schur's lemma[8] it follows that $A_{ij}(s) = 0$ unless $\mathfrak{A}_H^{(i)}$ is equivalent to $\mathfrak{B}_H^{(j)}$. For a fixed index j there can be only one index i such that this is so, since $\mathfrak{A}_H^{(1)}, \cdots, \mathfrak{A}_H^{(m)}$ are all inequivalent. Hence there can be only one non-vanishing component-matrix $A_{ij}(s)$ in each column of $A(s)$, which means of course that $A(s)$ merely permutes the spaces $\mathfrak{R}_1, \cdots, \mathfrak{R}_m$ among each other.

Since $\mathfrak{A}_G$ is irreducible it is transitive, and all the spaces $\mathfrak{R}_i$ must have the same dimension. Since by Theorem 1 all the irreducible components of $\mathfrak{A}_H$ have the same degree n, it follows that $\mathfrak{A}_H^{(1)}, \cdots, \mathfrak{A}_H^{(m)}$ must each occur in $\mathfrak{A}_H$ the same number of times l. That the degree of $\mathfrak{A}_G$ is lmn is obvious.

We may give a more explicit form for $\mathfrak{A}_G$ by introducing the group G' leaving the first space $\mathfrak{R}_1$ invariant. G' consists of all elements s' of G for which $s'\mathfrak{R}_1 = \mathfrak{R}_1$, or in other words such that $A_{11}(s')$ is the non-vanishing component-matrix in the first row and column of $A(s')$. If we select elements $r_2, \cdots, r_m$ in G such that $r_i\mathfrak{R}_1 = \mathfrak{R}_i$ then

$$G = G' + r_2G' + \cdots + r_mG'.$$

For each s in G throws $\mathfrak{R}_1$ into some $\mathfrak{R}_i$, whence $s^{-1}r_i$ is in G'. G' is thus a (not in general invariant) subgroup of G of finite index m.

---

[8] If $\mathfrak{A}_H$ and $\mathfrak{B}_H$ are multiples of inequivalent irreducible representations of a group H, and $A(u)C = CB(u)$ for all u in H, then $C = 0$; this is valid in any number field.

The non-vanishing component-matrix in the first column of $A(r_i)$ is $A_{i1}(r_i)$. We can choose a basis in each of the spaces $\mathfrak{R}_2, \cdots, \mathfrak{R}_m$ such that $A_{i1}(r_i)$ is the identity matrix E of degree ln. This is effected by the transform $L^{-1}A(s)L$, where

$$L = \begin{Vmatrix} E & & & & \\ & A_{21}(r_2) & & & \\ & & \cdot & & \\ & & & \cdot & \\ & & & & A_{m1}(r_m) \end{Vmatrix}.$$

From $A_{i1}(r_i) = E$ we conclude at once that $A_{1i}(r_i^{-1}) = E$, and hence

$$A_{ij}(r_i s' r_j^{-1}) = \sum_{k,l} A_{ik}(r_i) A_{kl}(s') A_{lj}(r_j^{-1}) = A_{11}(s')$$

for all s' in G', since $k = 1$ and $l = 1$ are the only indices for which $A_{ik}(r_i)$ and $A_{lj}(r_j^{-1})$ do not vanish.

For a given s in G and a given index j, there is a uniquely determined index i such that $r_i^{-1} s r_j$ is in G', namely that of the coset $r_i G'$ to which sr_j belongs. Setting $r_i^{-1} s r_j = s'$, then $s = r_i s' r_j^{-1}$ and

$$A_{ij}(s) = A_{11}(r_i^{-1} s r_j).$$

On writing $A'(s')$ in place of $A_{11}(s')$, the component-matrices $A_{ij}(s)$ of $A(s)$ for any s in G are given by

$$A_{ij}(s) = \begin{cases} A'(r_i^{-1} s r_j) & \text{if } r_i^{-1}\, s r_j \text{ is in } G', \\ 0 & \text{otherwise.} \end{cases}$$

From this we see that the representation $\mathfrak{A}'_{G'}$ of G' defined by $s' \to A'(s')$ must be irreducible, for its reducibility would evidently entail that of $\mathfrak{A}_G$. $\mathfrak{A}_G$ is the imprimitive representation of G *generated* by the irreducible representation $\mathfrak{A}'_{G'}$ of G'. We add the final remark that $\mathfrak{A}'_H = l \cdot \mathfrak{A}_H^{(1)}$; the representation $\mathfrak{A}'_H$ induced by $\mathfrak{A}'_{G'}$ in the invariant subgroup H of G' is simply the multiple l of the irreducible $\mathfrak{A}_H^{(1)}$.

3. Structure of $\mathfrak{A}'_{G'}$.

Since $\mathfrak{A}_G$ is completely known when we know $\mathfrak{A}'_{G'}$ we may proceed to investigate the structure of the latter. *In this section we shall assume that the groundfield* P *is algebraically closed.* Since we shall now operate entirely within G' we may avoid a multitude of primes by replacing G' by G, $\mathfrak{A}'_{G'}$ by $\mathfrak{A}_G$. We deal now with the case $m = 1$, when all the irreducible components of $\mathfrak{A}_H$ are equivalent: $\mathfrak{A}_H = l \cdot \mathfrak{A}_H^{(1)}$. Again denoting by n the degree of $\mathfrak{A}_H^{(1)}$, that of $\mathfrak{A}_G$ is ln.

Theorem 3. *When all the irreducible components of $\mathfrak{A}_H$ are equivalent,*

$\mathfrak{A}_H = l \cdot \mathfrak{A}_H^{(1)}$, *and the ground-field* P *is algebraically closed, then* $\mathfrak{A}_G$ *is the direct product of two irreducible projective representations of* G:

$$A(s) = C(s) \times \Gamma(s),$$

where $\Gamma(s)$ *is of degree* l, *and* $C(s)$ *has the same degree* n *as* $\mathfrak{A}_H^{(1)}$. $s \to \Gamma(s)$ *is actually a projective representation of the factor-group* G/H.

PROOF: We choose a coordinate system in $\mathfrak{R}$ such that the matrices $A(u)$ of $\mathfrak{A}_H$ have the fully reduced form

$$A(u) = \begin{Vmatrix} A_1(u) & & & & \\ & A_1(u) & & & \\ & & \cdot & & \\ & & & \cdot & \\ & & & & A_1(u) \end{Vmatrix},$$

and write correspondingly[9]

$$A(s) = \begin{Vmatrix} A_{11}(s) & \cdots & A_{1l}(s) \\ \cdots & \cdots & \cdots \\ A_{l1}(s) & \cdots & A_{ll}(s) \end{Vmatrix}.$$

Each component $A_{\alpha\beta}(s)$ of $A(s)$ is itself a matrix of degree n. For elements u in H, $A_{\alpha\beta}(u)$ is zero if $\alpha \neq \beta$ and $A_{\alpha\alpha}(u) = A_1(u)$.

By Theorem 1, every conjugate of $\mathfrak{A}_H^{(1)}$ relative to G is equivalent to $\mathfrak{A}_H^{(1)}$. Hence to each s in G there corresponds a non-singular matrix $C(s)$ of degree n such that

$$A_1(s^{-1}us) = C^{-1}(s)A_1(u)C(s)$$

for all u in H. From

$$A(u)A(s) = A(s)A(s^{-1}us)$$

we have at once

$$A_1(u)A_{\alpha\beta}(s) = A_{\alpha\beta}(s)A_1(s^{-1}us),$$

and consequently

$$A_1(u)A_{\alpha\beta}(s)C^{-1}(s) = A_{\alpha\beta}(s)C^{-1}(s)A_1(u).$$

Regarding s as fixed, this equation holds for all u in H, and hence by the second part of Schur's lemma,[10] $A_{\alpha\beta}(s)C^{-1}(s)$ is a scalar matrix:

$$A_{\alpha\beta}(s) = \gamma_{\alpha\beta}(s)C(s),$$

[9] $A_{11}(s)$ should not be confused with the $A_{11}(s)$ in §2; it really stands for $A_{11}'(s')$. In the case we are considering now, the developments of §2 are of course entirely absent.

[10] In an algebraically closed field, the only matrices commuting with an irreducible set of matrices are scalar multiples of the identity matrix.

The non-vanishing component-matrix in the first column of $A(r_i)$ is $A_{i1}(r_i)$. We can choose a basis in each of the spaces $\mathfrak{R}_2, \cdots, \mathfrak{R}_m$ such that $A_{i1}(r_i)$ is the identity matrix E of degree ln. This is effected by the transform $L^{-1}A(s)L$, where

$$L = \begin{Vmatrix} E & & & & \\ & A_{21}(r_2) & & & \\ & & \cdot & & \\ & & & \cdot & \\ & & & & A_{m1}(r_m) \end{Vmatrix}.$$

From $A_{i1}(r_i) = E$ we conclude at once that $A_{1i}(r_i^{-1}) = E$, and hence

$$A_{ij}(r_i s' r_j^{-1}) = \sum_{k,l} A_{ik}(r_i)A_{kl}(s')A_{lj}(r_j^{-1}) = A_{11}(s')$$

for all s' in G', since $k = 1$ and $l = 1$ are the only indices for which $A_{ik}(r_i)$ and $A_{lj}(r_j^{-1})$ do not vanish.

For a given s in G and a given index j, there is a uniquely determined index i such that $r_i^{-1}sr_j$ is in G', namely that of the coset r_iG' to which sr_j belongs. Setting $r_i^{-1}sr_j = s'$, then $s = r_is'r_j^{-1}$ and

$$A_{ij}(s) = A_{11}(r_i^{-1}sr_j).$$

On writing $A'(s')$ in place of $A_{11}(s')$, the component-matrices $A_{ij}(s)$ of $A(s)$ for any s in G are given by

$$A_{ij}(s) = \begin{cases} A'(r_i^{-1}sr_j) & \text{if } r_i^{-1}sr_j \text{ is in } G', \\ 0 & \text{otherwise.} \end{cases}$$

From this we see that the representation $\mathfrak{A}'_{G'}$ of G' defined by $s' \to A'(s')$ must be irreducible, for its reducibility would evidently entail that of $\mathfrak{A}_G$. $\mathfrak{A}_G$ is the imprimitive representation of G *generated* by the irreducible representation $\mathfrak{A}'_{G'}$ of G'. We add the final remark that $\mathfrak{A}'_H = l \cdot \mathfrak{A}_H^{(1)}$; the representation $\mathfrak{A}'_H$ induced by $\mathfrak{A}'_{G'}$ in the invariant subgroup H of G' is simply the multiple l of the irreducible $\mathfrak{A}_H^{(1)}$.

3. Structure of $\mathfrak{A}'_{G'}$.

Since $\mathfrak{A}_G$ is completely known when we know $\mathfrak{A}'_{G'}$ we may proceed to investigate the structure of the latter. *In this section we shall assume that the groundfield* P *is algebraically closed.* Since we shall now operate entirely within G' we may avoid a multitude of primes by replacing G' by G, $\mathfrak{A}'_{G'}$ by $\mathfrak{A}_G$. We deal now with the case $m = 1$, when all the irreducible components of $\mathfrak{A}_H$ are equivalent: $\mathfrak{A}_H = l \cdot \mathfrak{A}_H^{(1)}$. Again denoting by n the degree of $\mathfrak{A}_H^{(1)}$, that of $\mathfrak{A}_G$ is ln.

THEOREM 3. *When all the irreducible components of* $\mathfrak{A}_H$ *are equivalent,*

$\mathfrak{A}_H = l \cdot \mathfrak{A}_H^{(1)}$, *and the ground-field* P *is algebraically closed, then* $\mathfrak{A}_G$ *is the direct product of two irreducible projective representations of* G:

$$A(s) = C(s) \times \Gamma(s),$$

where $\Gamma(s)$ *is of degree* l, *and* $C(s)$ *has the same degree* n *as* $\mathfrak{A}_H^{(1)}$. $s \to \Gamma(s)$ *is actually a projective representation of the factor-group* G/H.

PROOF: We choose a coordinate system in $\mathfrak{R}$ such that the matrices $A(u)$ of $\mathfrak{A}_H$ have the fully reduced form

$$A(u) = \left\| \begin{matrix} A_1(u) & & & \\ & A_1(u) & & \\ & & \ddots & \\ & & & A_1(u) \end{matrix} \right\|,$$

and write correspondingly[9]

$$A(s) = \left\| \begin{matrix} A_{11}(s) & \cdots & A_{1l}(s) \\ \cdots & \cdots & \cdots \\ A_{l1}(s) & \cdots & A_{ll}(s) \end{matrix} \right\|.$$

Each component $A_{\alpha\beta}(s)$ of $A(s)$ is itself a matrix of degree n. For elements u in H, $A_{\alpha\beta}(u)$ is zero if $\alpha \neq \beta$ and $A_{\alpha\alpha}(u) = A_1(u)$.

By Theorem 1, every conjugate of $\mathfrak{A}_H^{(1)}$ relative to G is equivalent to $\mathfrak{A}_H^{(1)}$. Hence to each s in G there corresponds a non-singular matrix $C(s)$ of degree n such that

$$A_1(s^{-1}us) = C^{-1}(s)A_1(u)C(s)$$

for all u in H. From

$$A(u)A(s) = A(s)A(s^{-1}us)$$

we have at once

$$A_1(u)A_{\alpha\beta}(s) = A_{\alpha\beta}(s)A_1(s^{-1}us),$$

and consequently

$$A_1(u)A_{\alpha\beta}(s)C^{-1}(s) = A_{\alpha\beta}(s)C^{-1}(s)A_1(u).$$

Regarding s as fixed, this equation holds for all u in H, and hence by the second part of Schur's lemma,[10] $A_{\alpha\beta}(s)C^{-1}(s)$ is a scalar matrix:

$$A_{\alpha\beta}(s) = \gamma_{\alpha\beta}(s)C(s),$$

---

[9] $A_{11}(s)$ should not be confused with the $A_{11}(s)$ in §2; it really stands for $A'_{11}(s')$. In the case we are considering now, the developments of §2 are of course entirely absent.

[10] In an algebraically closed field, the only matrices commuting with an irreducible set of matrices are scalar multiples of the identity matrix.

where $\gamma_{\alpha\beta}(s)$ is in P. This means of course that the matrix $A(s)$ is the direct (Kronecker) product

$$A(s) = C(s) \times \Gamma(s),$$

where $\Gamma(s)$ is the matrix $\| \gamma_{\alpha\beta}(s) \|$ of degree l. By Schur's lemma, $C(s)$ is determined by s to within an arbitrary scalar multiple; if $C(s)$ be replaced by $\rho C(s)$ then $\Gamma(s)$ must of course be replaced by $\rho^{-1}\Gamma(s)$.

If s and t are any two elements of G, then from the two equations

$$A(st) = C(st) \times \Gamma(st)$$

$$A(s)A(t) = C(s)C(t) \times \Gamma(s)\Gamma(t)$$

it follows that $C(st)$ and $\Gamma(st)$ differ from $C(s)C(t)$ and $\Gamma(s)\Gamma(t)$ respectively only by scalar multiples. $s \to C(s)$ and $s \to \Gamma(s)$ are therefore *projective* representations of G. If either were reducible, the same would be true of $\mathfrak{A}_G$.

Since the matrices $A(u)$ of $\mathfrak{A}_H$ have the form

$$A(u) = A_1(u) \times E_l,$$

E_l being the identity matrix of degree l, we can assume $\Gamma(u) = E_l$ for all u in H. Hence $s \to \Gamma(s)$ is actually a projective representation of G/H.

4. The Embedding Problem

We turn next to the problem of embedding a given irreducible representation $\mathfrak{A}_H^{(1)}$ of H in an irreducible representation $\mathfrak{A}_G$ of G: *can we find an irreducible $\mathfrak{A}_G$ such that $\mathfrak{A}_H^{(1)}$ occurs as an irreducible component in $\mathfrak{A}_H$?*

An obviously necessary condition is that $\mathfrak{A}_H^{(1)}$ have a finite number of inequivalent conjugates relative to G. If we define G' to be the group leaving $\mathfrak{A}_H^{(1)}$ invariant, consisting of all elements s' in G such that the conjugate $u \to A_1(s'^{-1}us')$ of $\mathfrak{A}_H^{(1)}$ is equivalent to $\mathfrak{A}_H^{(1)}$, then this condition evidently means that G' shall be of finite index in G. This G' is the same as that introduced in §2, and hence a second obviously necessary condition is that it be possible to embed $\mathfrak{A}_H^{(1)}$ in an irreducible representation $\mathfrak{A}'_{G'}$ of G'.

But these conditions are also sufficient. Let m be the number of distinct conjugates of $\mathfrak{A}_H^{(1)}$ relative to G. Then there exist m elements $r_1(=1), r_2, \cdots, r_m$ in G such that

$$\mathfrak{A}_H^{(i)}: u \to A_i(u) = A_1(r_i^{-1}ur_i), \qquad (i = 1, 2, \cdots, m),$$

are the m distinct conjugates of $\mathfrak{A}_H^{(1)}$, and

$$G = r_1G' + r_2G' + \cdots + r_mG'.$$

By hypothesis there exists an irreducible representation $\mathfrak{A}'_{G'}$ of G', $s' \to A'(s')$, such that $\mathfrak{A}_H^{(1)}$ is an irreducible component of the representation $\mathfrak{A}'_H$ of H induced in H by $\mathfrak{A}'_{G'}$. Since $\mathfrak{A}_H^{(1)}$ is equivalent to all its conjugates relative to G', $\mathfrak{A}'_H$ is necessarily a multiple l of $\mathfrak{A}_H^{(1)}$.

We now define $\mathfrak{A}_G$ to be the representation of G generated by $\mathfrak{A}'_{G'}$, that is, we define

$$A(s) = \begin{Vmatrix} A_{11}(s) & \cdots & A_{1m}(s) \\ \cdots & \cdots & \cdots \\ A_{m1}(s) & \cdots & A_{mm}(s) \end{Vmatrix}$$

where

$$A_{ij}(s) = \begin{cases} A'(r_i^{-1}sr_j) & \text{if } r_i^{-1}sr_j \text{ is in } G', \\ 0 & \text{otherwise.} \end{cases}$$

To see that $s \to A(s)$ is in fact a representation $\mathfrak{A}_G$ of G, we note that $\sum_j A_{ij}(s)A_{jk}(t)$ will be zero unless there is an index j such that $r_i^{-1}sr_j$ and $r_j^{-1}tr_k$ are both in G', and in this case it is equal to

$$A'(r_i^{-1}sr_j)A'(r_j^{-1}tr_k) = A'(r_i^{-1}str_k) = A_{ik}(st).$$

Hence

$$A(s)A(t) = A(st).$$

For elements u in H, $A_{ij}(u) = 0$ if $i \neq j$, and

$$A_{ii}(u) = A'(r_i^{-1}ur_i) = \begin{Vmatrix} A_i(u) & & & & \\ & A_i(u) & & & \\ & & \cdot & & \\ & & & \cdot & \\ & & & & \cdot \\ & & & & A_i(u) \end{Vmatrix}.$$

Hence

$$\mathfrak{A}_H = l \cdot \mathfrak{A}_H^{(1)} + l \cdot \mathfrak{A}_H^{(2)} + \cdots + l \cdot \mathfrak{A}_H^{(m)}.$$

To see that $\mathfrak{A}_G$ is irreducible, let $\mathfrak{B}_G$ be the irreducible component of $\mathfrak{A}_G$ such that $\mathfrak{B}_{G'}$ contains $\mathfrak{A}'_{G'}$ (this does not presume the full reducibility of $\mathfrak{A}_G$!) But then $\mathfrak{B}_H$ will contain $\mathfrak{A}'_H = l \cdot \mathfrak{A}_H^{(1)}$ and hence by Theorem 1 will contain $l \cdot \mathfrak{A}_H^{(2)}, \cdots, l \cdot \mathfrak{A}_H^{(m)}$ as well. $\mathfrak{B}_G$ has therefore the same degree as $\mathfrak{A}_G$, and so must coincide with it.

This reduces the problem to that of embedding $\mathfrak{A}_H^{(1)}$ in an irreducible representation $\mathfrak{A}'_{G'}$ of G'. While the foregoing was valid in any ground-field P, we now assume P to be algebraically closed. To simplify the notation we may again replace G' by G and assume that the given irreducible representation $\mathfrak{A}_H^{(1)}$ of H is equivalent to all its conjugates under G. This means (as in §3) that to each s in G there corresponds a non-singular matrix $C(s)$, of the same degree n as the matrices $A_1(u)$ of $\mathfrak{A}_H^{(1)}$, such that

$$A_1(s^{-1}us) = C^{-1}(s)A_1(u)C(s)$$

for all u in H. From Schur's lemma we readily find that $C(s)$ is determined by s to within an arbitrary scalar multiple, and that $C(s)C(t)$ differs from $C(st)$ only by a scalar multiple:

$$C(s)C(t) = \alpha(s, t)C(st),$$

with $\alpha(s, t) \neq 0$ in P. $s \to C(s)$ is of course the same projective representation of degree n of G that occurred in §3; the essential point is that it is determined entirely by $\mathfrak{A}_H^{(1)}$ and the structure of H in G.

The function $\alpha(s, t)$ must satisfy the associative condition

$$\alpha(rs, t)\alpha(r, s) = \alpha(r, st)\alpha(s, t),$$

and such a function is called a *factor-set*.[11] If we make a new choice of the representative matrices $C(s)$, say $\sigma(s)C(s)$, then the factor-set $\alpha(s, t)$ is replaced by the *associate* factor-set

$$\frac{\sigma(s)\,\sigma(t)}{\sigma(st)}\,\alpha(s, t).$$

The reciprocals $\alpha^{-1}(s, t)$ of the numbers $\alpha(s, t)$ also satisfy the associative condition; $\alpha^{-1}(s, t)$ is the factor-set *inverse* to $\alpha(s, t)$.

We proceed to show that the factor-set $\alpha(s, t)$ belonging to the projective representation $s \to C(s)$ of G is associate to a factor-set essentially of G/H. We need only choose representatives $s_0, t_0, \cdots$ from each coset $s_0H, t_0H, \cdots$ of G mod H, fix $C(s_0), C(t_0), \cdots$ in any way, and define

$$C(s_0u) = C(s_0)A_1(u), \qquad C(u) = A_1(u).$$

Then

$$C(su) = C(s)A_1(u)$$

for all s in G and all u in H. Hence if u and v are in H, s and t in G,

$$\begin{aligned} C(su)C(tv) &= C(s)A_1(u)C(t)A_1(v) \\ &= C(s)C(t)A_1(t^{-1}ut)A_1(v) \\ &= \alpha(s, t)C(st)A_1(t^{-1}utv) \\ &= \alpha(s, t)C(sutv), \end{aligned}$$

whence

$$\alpha(su, tv) = \alpha(s, t).$$

Thus the values of the function $\alpha(s, t)$ depend only on the cosets of G mod H in which s and t lie.

Now in Theorem 3 the matrices $\Gamma(s)$ must evidently multiply according to the inverse factor-set $\alpha^{-1}(s, t)$:

$$\Gamma(s)\Gamma(t) = \alpha^{-1}(s, t)\Gamma(st).$$

---

[11] See, for example, van der Waerden, l.c. §21, and the references there to Schur's original papers.

Consequently a necessary condition that $\mathfrak{A}_H^{(1)}$ can be embedded in an irreducible $\mathfrak{A}_G$ is that the factor-set $\alpha^{-1}(s, t)$ of G/H can be realized by a projective representation $s \to \Gamma(s)$ of G/H.

But this condition is also sufficient. If such a realization exists then we may take any irreducible component of it, call it $s \to \Gamma(s)$, and define

$$A(s) = C(s) \times \Gamma(s).$$

$s \to A(s)$ is evidently an ordinary representation $\mathfrak{A}_G$ of G, and $\mathfrak{A}_H = l \cdot \mathfrak{A}_H^{(1)}$ if l is the degree of $\Gamma(s)$. To show the irreducibility of $\mathfrak{A}_G$, suppose that we have a linear relation among the components of $A(s)$:

$$\sum_{i,j,\alpha,\beta} \rho_{ij\alpha\beta} c_{ij}(s) \gamma_{\alpha\beta}(s) = 0, \qquad \begin{pmatrix} i, j = 1, 2, \cdots, n \\ \alpha, \beta = 1, 2, \cdots, l \end{pmatrix},$$

where

$$C(s) = \| c_{ij}(s) \|, \qquad \Gamma(s) = \| \gamma_{\alpha\beta}(s) \|.$$

Let s_0 be a fixed element of G, and u a variable element of H. Since $\Gamma(s_0u) = \Gamma(s_0)$ we have

$$\sum \rho_{ij\alpha\beta} c_{ij}(s_0u) \gamma_{\alpha\beta}(s_0) = 0.$$

Now $C(s_0u) = C(s_0)A_1(u)$, and since there are n^2 linearly independent matrices $A_1(u)$ as u ranges over H (by Burnside's theorem), it follows that there are n^2 linearly independent matrices $C(s_0u)$. Hence

$$\sum_{\alpha,\beta} \rho_{ij\alpha\beta} \gamma_{\alpha\beta}(s_0) = 0. \qquad (i, j = 1, 2, \cdots, n).$$

This holds for each s_0 in G, and since $s \to \Gamma(s)$ is irreducible we conclude that every $\rho_{ij\alpha\beta} = 0$.

We may give an alternative expression of the condition that $\alpha^{-1}(s, t)$ be realizable by a projective representation of G/H by introducing the group-algebra $\mathfrak{a}$ of G/H relative to the factor-set $\alpha^{-1}(s, t)$. Let us denote the elements of G/H by $S, T, \cdots$; then we may write

$$\alpha_{S,T} = \alpha(s, t)$$

if s belongs to the coset S and t to T. Corresponding to each S in G/H we introduce a symbol w_S and define $\mathfrak{a}$ to be the algebra over P with the symbols w_S as a basis, the product $w_S w_T$ of the two basic elements w_S and w_T being defined by

$$w_S w_T = \alpha_{S,T}^{-1} w_{ST} .$$

$\mathfrak{a}$ is an associative algebra because of the associativity conditions on $\alpha(s, t)$. Its order over P is the order of G/H, which may of course be infinite.

Any representation of $\mathfrak{a}$ leads at once to a projective representation of G/H which realizes the factor-set $\alpha_{S,T}^{-1}$, and conversely. The condition in question is therefore equivalent to the condition that the algebra $\mathfrak{a}$ possess a representa-

tion of finite degree. This is certainly the case if H is of finite index in G, for then $\mathfrak{a}$ is an algebra of finite order over P.

We summarize the results of this section in the following:

Theorem 4. *A given irreducible representation $\mathfrak{A}_H^{(1)}$ of H can be embedded in an irreducible representation $\mathfrak{A}_G$ of G if and only if* (i) *the subgroup G' of G leaving $\mathfrak{A}_H^{(1)}$ invariant is of finite index in G, i.e. the number of inequivalent conjugates of $\mathfrak{A}_H^{(1)}$ relative to G is finite, and* (ii) *$\mathfrak{A}_H^{(1)}$ can be embedded in an irreducible representation $\mathfrak{A}'_{G'}$ of G'.*

If the ground-field P is algebraically closed, (ii) *holds if and only if* (iii) *the factor-set $\alpha^{-1}(s', t')$ of G'/H determined by $\mathfrak{A}_H^{(1)}$ can be realized by a projective representation $s' \to \Gamma(s')$ of G'/H, or in other words if the group-algebra $\mathfrak{a}$ of G'/H corresponding to the factor-set $\alpha^{-1}(s', t')$ has a representation of finite degree.*

In particular, if H is of finite index in G, then every irreducible representation of H can be embedded in an irreducible representation of G.

The last assertion can be proved more simply, and without any assumptions on the ground-field P, by forming the representation $\bar{\mathfrak{A}}_G$ of G generated by $\mathfrak{A}_H^{(1)}$. Suppose

$$G = t_1H + t_2H + \cdots + t_kH, \qquad (t_1 = 1).$$

Then $\bar{\mathfrak{A}}_G$ is defined as usual by

$$\bar{A}(s) = \left\| \begin{matrix} \bar{A}_{11}(s) & \cdots & \bar{A}_{1k}(s) \\ \cdots & \cdots & \cdots \\ \bar{A}_{k1}(s) & \cdots & \bar{A}_{kk}(s) \end{matrix} \right\|$$

where

$$\bar{A}_{ij}(s) = \begin{cases} A_1(t_i^{-1}st_j) & \text{if } t_i^{-1}st_j \text{ is in } H, \\ 0 & \text{otherwise.} \end{cases}$$

If m is the index of G' in G, and h that of H in $G'(mh = k)$, then evidently

$$\bar{\mathfrak{A}}_H = h\cdot\mathfrak{A}_H^{(1)} + h\cdot\mathfrak{A}_H^{(2)} + \cdots + h\cdot\mathfrak{A}_H^{(m)}$$

where $\mathfrak{A}_H^{(1)}, \mathfrak{A}_H^{(2)}, \cdots, \mathfrak{A}_H^{(m)}$ are the m distinct conjugates of $\mathfrak{A}_H^{(1)}$. Hence if $\mathfrak{A}_G$ is any irreducible constituent of $\bar{\mathfrak{A}}_G$ (again we do not assume the full reducibility of $\bar{\mathfrak{A}}_G$), then the irreducible components of $\mathfrak{A}_H$ must be $\mathfrak{A}_H^{(1)}$ and its conjugates.

5. Associates

We consider finally the question of finding all possible ways of embedding an irreducible representation of H in an irreducible representation of G. Two irreducible representations $\mathfrak{A}_G$ and $\mathfrak{B}_G$ of G will be called *associate*[12] relative to H if $\mathfrak{A}_H$ and $\mathfrak{B}_H$ have an irreducible component in common. If we denote by $\mathfrak{A}_H^{(1)}$ this common irreducible component, then, by Theorem 1, $\mathfrak{A}_H$ and $\mathfrak{B}_H$ have

[12] This term is perhaps unfortunate, as it suggests that $\mathfrak{A}_G$ and $\mathfrak{B}_G$ are projectively equivalent; this turns out to be the case if G/H is a finite cyclic group, but is not so in general.

the same irreducible components, namely the conjugates of $\mathfrak{A}_H^{(1)}$ relative to G, except possibly for their multiplicity.

Now the group G', the irreducible projective representation $s' \to C(s')$ of G' whose degree n is the same as that of $\mathfrak{A}_H^{(1)}$, and the corresponding factor-set $\alpha(s', t')$ of G'/H are all determined by $\mathfrak{A}_H^{(1)}$. We are then allowed the liberty of picking any irreducible projective representation $s' \to \Gamma(s')$ of G'/H which realizes the inverse factor-set $\alpha^{-1}(s', t')$. Having done this, the resulting irreducible representation $\mathfrak{A}'_{G'}$ of G' uniquely determines the irreducible representation $\mathfrak{A}_G$ of G, and conversely $\mathfrak{A}'_{G'}$ is uniquely determined by $\mathfrak{A}_G$ since it can be characterized as that irreducible component of $\mathfrak{A}_{G'}$ which induces a multiple of $\mathfrak{A}_H^{(1)}$ in H. Of the three stages in this construction, the first and last are completely fixed; two associates $\mathfrak{A}_G$ and $\mathfrak{B}_G$ can therefore differ only in the middle stage, the choice of Γ.

Let us fix upon a definite determination of the matrices $C(s')$, and let Γ and Δ be the projective representations of G'/H corresponding to the associates $\mathfrak{A}_G$ and $\mathfrak{B}_G$ respectively, so that

$$A'(s') = C(s') \times \Gamma(s')$$
$$B'(s') = C(s') \times \Delta(s').$$

Since the matrices $C(s')$ are fixed, so also are $\Gamma(s')$ and $\Delta(s')$. We shall say that Γ and Δ are *strictly*[13] equivalent if and only if there exists a constant non-singular matrix N such that

$$\Delta(s') = N^{-1}\Gamma(s')N$$

for all s' in G'. This amounts to saying that the corresponding representations of the algebra $\mathfrak{a}$ are equivalent in the ordinary sense.

If such an N exists, then $\mathfrak{A}'_{G'}$ and $\mathfrak{B}'_{G'}$ are obviously equivalent. If, on the other hand, $\mathfrak{A}'_{G'}$ and $\mathfrak{B}'_{G'}$ are equivalent, then there exists a constant non-singular matrix M such that

$$C(s') \times \Delta(s') = M^{-1}[C(s') \times \Gamma(s')]M.$$

Γ and Δ must evidently have the same degree l. For elements u in H this becomes

$$A_1(u) \times E_l = M^{-1}[A_1(u) \times E_l]M.$$

This means that M commutes with the matrices

$$\begin{Vmatrix} A_1(u) & & & & \\ & A_1(u) & & & \\ & & \cdot & & \\ & & & \cdot & \\ & & & & A_1(u) \end{Vmatrix}$$

[13] The purpose of this adjective is to prevent possible confusion with *projective* equivalence of Γ and Δ.

of $l \cdot \mathfrak{A}_H^{(1)}$, and so must be of the form $E_n \times N$, N being of course a non-singular matrix of degree l. We then have

$$C(s') \times \Delta(s') = C(s') \times N^{-1}\Gamma(s')N.$$

If s_0' is a fixed element of G' then by Burnside's theorem there are n^2 linearly independent matrices $C(s_0' u)$ as u ranges over H, from which we readily conclude that

$$\Delta(s') = N^{-1}\Gamma(s')N.$$

Since $\mathfrak{A}_G$ and $\mathfrak{B}_G$ are equivalent if and only if $\mathfrak{A}_{G'}'$ and $\mathfrak{B}_{G'}'$ are equivalent, we have:

THEOREM 5. *Two associates of G relative to H can differ only in the projective representations of G'/H which they determine, and they are equivalent if and only if the latter are strictly equivalent.*

We can visualize the relationship between associates of G relative to H on the one hand, and conjugates of H relative to G on the other, by dissecting the set of all irreducible representations of G into classes of associates relative to H, and similarly the set of all irreducible representations of H into classes of conjugates relative to G. The latter fall into two categories: (I) those that can be embedded, and (II) those that cannot. We then have a one-to-one correspondence between the class of classes of associates of G and the class of those classes of conjugates of H which fall in Category I. The relationship between a class of associates and the corresponding class of conjugates is simply that each member of the former induces in H a representation whose irreducible components are precisely all the members of the latter.

When H is of finite index in G, Category II is empty, and moreover each class of associates is finite, since there is a one-to-one correspondence between it and the class of all the inequivalent irreducible representations of the algebra $\mathfrak{a}$. In this case we may say that if we know all the irreducible representations of G then we know all those of H (the converse being true, at least theoretically, even when H is not of finite index in G). Every irreducible representation of H can be found by decomposing the representations induced by those of G, and every irreducible representation of G can be found by reducing the representations generated by those of H.

In fact, the irreducible constituents of the representation $\bar{\mathfrak{A}}_G$ generated by an irreducible representation $\mathfrak{A}_H^{(1)}$ of H (end of §4) comprise the whole class of corresponding asssociates. Let us denote the latter by ${}^{(1)}\mathfrak{A}_G$, ${}^{(2)}\mathfrak{A}_G$, $\cdots$, ${}^{(q)}\mathfrak{A}_G$, and let $l_1, l_2, \cdots, l_q$ be their respective indices of multiplicity, that is:

$${}^{(\alpha)}\mathfrak{A}_H = l_\alpha \cdot \mathfrak{A}_H^{(1)} + l_\alpha \cdot \mathfrak{A}_H^{(2)} + \cdots + l_\alpha \cdot \mathfrak{A}_H^{(m)}, \qquad (\alpha = 1, 2, \cdots, q).$$

Then we can show, at any rate if the characteristic of P is zero, that $\bar{\mathfrak{A}}_G$ contains each ${}^{(\alpha)}\mathfrak{A}_G$ exactly l_α times, as in the Frobenius relations for finite groups. I am unable to say whether or not $\bar{\mathfrak{A}}_G$ is fully reducible.

The theory becomes especially simple when G/H is a finite cyclic group, say of order k. We may write

$$G = H + rH + r^2H + \cdots + r^{k-1}H,$$

where r^k, and no lower power of r, is in H. We shall assume here that k is not divisible by the characteristic of P. G'/H is also cyclic, and we know that every projective representation of a cyclic group can be normalized into an ordinary one (every factor-set is associate to 1). Every irreducible representation Γ of G'/H is, moreover, of degree one, and hence $l = 1$. The index m of G' in G is a divisor of k, and the number of distinct irreducible representations Γ of G'/H is equal to the order k/m of G'/H. *The irreducible components of the representation $\mathfrak{A}_H$ induced in H by an irreducible $\mathfrak{A}_G$ are all inequivalent, and their number m is a divisor of the index k of H in G; the number of distinct associates of $\mathfrak{A}_G$ relative to H is precisely k/m.*

On setting $q = k/m$ we may write

$$G = G' + rG' + r^2G' + \cdots + r^{m-1}G',$$

$$G' = H + r^mH + r^{2m}H + \cdots + r^{(q-1)m}H.$$

In order to emphasize the fact that the irreducible representation Γ of G'/H is of degree one, we shall write $\gamma(s')$ in place of $\Gamma(s')$. $\gamma(s')$ is, for each s' in G', a q^{th} root of unity, and its value depends only on the coset of G' mod H in which s' lies. Since moreover G'/H is a subgroup of the abelian group G/H,

$$\gamma(t^{-1}s't) = \gamma(s')$$

for every t in G and every s' in G'.

If $s \to B(s)$ is an associate $\mathfrak{B}_G$ of $\mathfrak{A}_G$, then, by Theorem 5, $\mathfrak{B}'_{G'}$ differs from $\mathfrak{A}'_{G'}$ only by such a one-dimensional representation of G'/H:

$$B'(s') = \gamma(s')A'(s').$$

On replacing r_i by r^{i-1} in the expression (§2) for the component-matrices $A_{ij}(s)$ of $A(s)$ we get

$$A_{ij}(s) = \begin{cases} A'(r^{-i+1}sr^{j-1}) & \text{if } r^{-i+1}sr^{j-1} \text{ is in } G', \\ 0 & \text{otherwise,} \end{cases}$$

and similarly of course for the components $B_{ij}(s)$ of $B(s)$. Since in the present case G' is itself invariant in G, we see that, for elements s' in G',

$$A_{ij}(s') = 0, \qquad (i \neq j),$$

$$A_{ii}(s') = A'(r^{-i+1}s'r^{i-1}),$$

Similarly,

$$B_{ij}(s') = 0, \qquad (i \neq j)$$
$$\begin{aligned} B_{ii}(s') &= B'(r^{-i+1}s'r^{i-1}) \\ &= \gamma(r^{-i+1}s'r^{i-1})A'(r^{-i+1}s'r^{i-1}) \\ &= \gamma(s')A_{ii}(s'). \end{aligned}$$

Hence

$$B(s') = \gamma(s')A(s').$$

Now $\gamma(r^m)$ is a q^{th} root of unity in P, and hence we can write

$$\gamma(r^m) = \epsilon^m$$

where ϵ is a k^{th} root of unity in P, recalling that $k = mq$. We assert that $\mathfrak{B}_G$ is equivalent to the representation $\mathfrak{A}_G^*$ of G defined by

$$A^*(r^i u) = \epsilon^i A(u), \qquad (i = 0, 1, \cdots, k-1; u \text{ in } H).$$

$\mathfrak{A}_G^*$ coincides with $\mathfrak{B}_G$ within G', i.e. $\mathfrak{A}_{G'}^* = \mathfrak{B}_{G'}$. But since G' is invariant in G, all the diagonal components $A_{ii}(s)$ of the matrix $A(s)$ corresponding to an element s of G not in G' must be zero. Hence $\mathfrak{A}_G^*$ and $\mathfrak{B}_G$ have the same character, and so must be equivalent.

This shows that every associate of $\mathfrak{A}_G$ is equivalent to an $\mathfrak{A}_G^*$ of the above form. But on the other hand every such $\mathfrak{A}_G^*$, defined by means of any k^{th} root of unity ϵ, is plainly associate to $\mathfrak{A}_G$, since $\mathfrak{A}_H^* = \mathfrak{A}_H$. This affords an easy means of writing down immediately all the associates of $\mathfrak{A}_G$. Expressed in this form, there are k possible associates of $\mathfrak{A}_G$ (corresponding to the k distinct k^{th} roots of unity ϵ in P), some of which may be equivalent, and similarly k possible conjugates of $\mathfrak{A}_H^{(1)}$ (corresponding to transforming by $1, r, r^2, \cdots r^{k-1}$). We then have a kind of reciprocity between them: if m be the number of inequivalent conjugates of $\mathfrak{A}_H^{(1)}$ relative to G, and q the number of inequivalent associates of $\mathfrak{A}_G$ relative to H, then the product mq of these numbers is equal to the index k of H in G. The case $k = 2$ was described in the introduction.

6. Extension to Semi-linear Representations

Let G be any group, P any field, and let there be a fixed homomorphic mapping of G onto a group of automorphisms of P. The effect on α in P of the automorphism corresponding to s in G will be denoted by α^s. In order to retain the privilege of writing operators on the left of operands it will be necessary to define the law of composition backwards:

$$\alpha^{st} = (\alpha^t)^s.$$

If $\mathfrak{R}$ is a vector-space over P of dimension n, and

$$x = (\xi_1, \xi_2, \cdots, \xi_n)$$

is a vector in R, then we define

$$x^s = (\xi_1^s, \xi_2^s, \cdots, \xi_n^s).$$

A *semi-linear representation* $\mathfrak{A}_G$ of G in $\mathfrak{R}$ is a homomorphic mapping $s \to \{A(s), s\}$ of G onto a group of semi-linear transformations

$$x \to A(s)x^s$$

in $\mathfrak{R}$. The result of $\{A(t), t\}$ followed by $\{A(s), s\}$ is

$$x \to A(s)[A(t)x^t]^s = A(s)A^s(t)x^{st},$$

and this must be the same as $\{A(st), st\}$, i.e.

$$x \to A(st)x^{st}.$$

Hence a necessary and sufficient condition that the correspondence $s \to \{A(s), s\}$ be a representation of G is that the matrices $A(s)$ satisfy

$$A(st) = A(s)A^s(t).$$

Passage to a new basis of $\mathfrak{R}$ connected with the old by a non-singular matrix M results in replacing $A(s)$ by $M^{-1}A(s)M^s$. Two semi-linear representations $\mathfrak{A}_G$ and $\mathfrak{B}_G$ of G are therefore said to be *equivalent* if their matrices $A(s)$ and $B(s)$ are connected by

$$B(s) = M^{-1}A(s)M^s.$$

A subspace $\mathfrak{S}$ of $\mathfrak{R}$ is *invariant* under $\mathfrak{A}_G$ if $A(s)x^s$ is in $\mathfrak{S}$ for all x in $\mathfrak{S}$ and all s in G. $\mathfrak{S}$ is *irreducible* if it contains no proper invariant subspace $\neq (0)$. $\mathfrak{A}_G$ is *fully reducible* if $\mathfrak{R}$ is the direct sum of irreducible invariant subspaces; this means that $A(s)$ decomposes into diagonal blocks $A_1(s)$, $A_2(s)$, $\cdots$ such that the semi-linear representations $s \to \{A_1(s), s\}$, $s \to \{A_2(s), s\}$, $\cdots$ are all irreducible.

Schur's lemma holds as well: *if $\mathfrak{A}_G$ and $\mathfrak{B}_G$ are irreducible, and if M is a constant matrix such that*

$$A(s)M^s = MB(s)$$

then $M = 0$ or M is non-singular. For the columns of M constitute an invariant subspace of the $\mathfrak{A}$-space, and the vectors x of the $\mathfrak{B}$-space satisfying $Mx = 0$ constitute an invariant subspace thereof. As an immediate corollary, such an M is zero if $\mathfrak{A}_G$ and $\mathfrak{B}_G$ are multiples of inequivalent semi-linear representations of G.

Let H be any invariant subgroup of G, and $\mathfrak{A}_H$ a semi-linear representation $u \to \{A(u), u\}$ of H. If s is a fixed element of G, then $u \to \{A^s(s^{-1}us), u\}$ is also a semi-linear representation of H which we call a *conjugate of* $\mathfrak{A}_H$ relative to G. For if we set

$$B(u) = A^s(s^{-1}us)$$

then

$$\begin{aligned} B(uv) &= A^s(s^{-1}us \cdot s^{-1}vs) \\ &= [A(s^{-1}us)A^{s^{-1}us}(s^{-1}vs)]^s \\ &= A^s(s^{-1}us)A^{us}(s^{-1}vs) = B(u)B^u(v). \end{aligned}$$

Notice that $u \to \{A(s^{-1}us), u\}$ without the superscript s is not in general a semi-linear representation of H.

The reasoning in §§1 and 2 can now be repeated without essential modification. Writing sx for $A(s)x^s$, we note that $s \cdot tx = st \cdot x$ and that $s \cdot \alpha x = \alpha^s \cdot sx$ (α in P). The only change that has to be made in the proof of Theorem 1 is in the expression for the effect of u in H operating on the basis $re_1, \cdots, re_n$ of $r\mathfrak{S}$:

$$\begin{aligned} u \cdot re_i &= r \cdot r^{-1}ur \cdot e_i \\ &= r \cdot \sum_j \alpha_{ij}(r^{-1}ur)e_j \\ &= \sum_j \alpha_{ij}^r(r^{-1}ur)re_j. \end{aligned}$$

This agrees with the definition of conjugate given above.

Proceeding as in §2, we obtain from

$$A(u)A^u(s) = A(s)A^s(s^{-1}us)$$

the equations

$$A_i(u)A_{ij}^u(s) = A_{ij}(s)A_j^s(s^{-1}us),$$

and apply Schur's lemma in its new form.

Transforming by the same constant matrix L, i.e. passing to $L^{-1}A(s)L^s$, we obtain, in the new coordinate system so defined, $A_{i1}(r_i) = E$. From

$$\sum_j A_{ij}(r_i)A_{ji}^{r_i}(r_i^{-1}) = A_{ii}(1) = E$$

we obtain $A_{1i}^{r_i}(r_i^{-1}) = E$, and hence $A_{1i}(r_i^{-1}) = E$.

From

$$A(rst) = A(r)A^r(s)A^{rs}(t)$$

we obtain

$$\begin{aligned} A_{ij}(r_i s' r_j^{-1}) &= \sum_{k,l} A_{ik}(r_i)A_{kl}^{r_i}(s')A_{lj}^{r_i s'}(r_j^{-1}) \\ &= A_{i1}(r_i)A_{11}^{r_i}(s')A_{1j}^{r_i s'}(r_j^{-1}) = A_{11}^{r_i}(s') \end{aligned}$$

for all s' in G'. Hence on setting $A_{11}(s') = A'(s')$ the expression for $\mathfrak{A}_G$ in terms of $\mathfrak{A}'_{G'}$ so defined is

$$A_{ij}(s) = \begin{cases} A'^{r_i}(r_i^{-1}sr_j) & \text{if } r_i^{-1}sr_j \text{ is in } G', \\ 0 & \text{otherwise.} \end{cases}$$

THE INSTITUTE FOR ADVANCED STUDY

Annals of Mathematics
Vol. 38, No. 3, July, 1937

ARITHMETIC IN FIELDS OF FORMAL POWER SERIES IN SEVERAL VARIABLES

By O. E. G. Schilling

(Received October 22, 1936)

The essential algebraic properties of perfect fields are determined by the structure of the field of residue classes belonging to the prime ideal of the field with respect to which it is perfect. This fact was familiar from many properties and theorems, and it has been shown directly by H. Hasse and F. K. Schmidt that the *type* of a field which is perfect with respect to a discrete archimedean valuation—that is to say the group of values is equal to the additive group of all integers—is generally given by the field of residue classes.[1] The fields of formal power series in one variable with coefficients in a given field belong to that class. These theories of Hensel can be generalized to fields upon which a valuation function is defined whose values form an arbitrary linearly ordered abelian group. In the present paper we shall investigate the algebraic and arithmetic properties of fields whose value group is isomorphic with the additive group of lexicographically ordered m-tuples of integers. Obviously reductions of such groups on archimedean groups will play an important rôle in our discussion. Nevertheless we have no general argument of induction that would apply throughout in the proofs, and the proofs though often similar have to be repeated to some extent. Under certain restricting conditions upon the group of values and the fields of residue classes we shall be able to show that our fields are isomorphic with fields of power series in several variables, furthermore that the valuations of the fields under consideration are uniquely determined. As in Hensel's case the arithmetic of these fields and that of their finite algebraic extensions are determined by the structure of the residue class field.

Furthermore we shall generalize the notion of ramification number which is associated with a branch point of an algebraic Riemann surface. The class of ramification meromorphisms here introduced describes the different types of ramifications that can occur on an abstract algebraic manifold. We hope that applications to classical problems of algebraic geometry can be found.

We are also interested in problems falling under the local class field theory. Under certain assumptions on the nature of the residue field we are in a position to develop certain analogs to the classical local class field theory. This leads to observations which, if properly analyzed, throw light on the relations between the different theorems of the known local class field theory. A more detailed investigation bearing on these questions will be published later. Finally we wish to remark that the normal finite algebras over fields of formal power

[1] H. Hasse & F. K. Schmidt, *Die Struktur diskret bewerteter Körper*, Crelle vol. 170 (1930).

series furnish within the frame of a general theory simple examples for division algebras which have different degree and exponent.

Fields of formal power series have been considered by T. Levi-Civita, H. Hahn, A. Ostrowski and W. Krull.[2] The latter introduced fields of formal power series for the solution of an existence theorem for perfect fields. The general theory of valuations as outlined by Krull was a natural tool for the following investigations.

1. Automorphisms and Meromorphisms of Discrete Value Groups of Finite Rank

Let $\Gamma = \Delta_m$ be a discrete linearly ordered abelian group of rank m; it is isomorphic with the additive group of linear forms $\sum_{j=1}^{m} a_j\alpha_j$ whose coefficients a_j lie in the additive group Γ^0 of all integers and which are lexicographically ordered from the right with respect to the subscript j. Its isolated subgroups Δ_i are given as the groups of all forms $\sum_{j=1}^{i} a_j\alpha_j$.[3] An automorphism of Γ is then an isomorphic mapping of Γ on itself such that the ordering of the group Γ is preserved. We now give a representation $\bar{\mathrm{A}}$ of the automorphism A of Γ with respect to a fixed basis $\alpha_1, \cdots, \alpha_m$ of Γ.

First we observe that $\Delta_i \times \mathrm{A} = \Delta_i$, because A conserves the ordering. The generators α_i of the factor groups Δ_i/Δ_{i-1} must be transformed into themselves modulo Δ_{i-1}. Therefore we obtain the following relations with integral coefficients a_{ij}:

$$\begin{aligned}
\alpha_1 \times \mathrm{A} &= \alpha_1 \\
\alpha_2 \times \mathrm{A} &= \alpha_1 a_{12} + \alpha_2 \\
&\cdots\cdots\cdots\cdots\cdots\cdots \\
\alpha_m \times \mathrm{A} &= \alpha_1 a_{1m} + \alpha_2 a_{2m} + \cdots + \alpha_{m-1} a_{m-1\,m} + \alpha_m ,
\end{aligned}$$

or written as a transformation

$$\Gamma = (\alpha_1, \cdots, \alpha_m) \times \mathrm{A} = (\alpha_1, \cdots, \alpha_m) \begin{Vmatrix} 1 & a_{12} & a_{13} & \cdots & a_{1m} \\ 0 & 1 & a_{23} & \cdots & \cdot \\ 0 & 0 & 1 & & \cdot \\ \cdot & \cdot & & \cdot & \cdot \\ \cdot & \cdot & & \cdot & \cdot \\ \cdot & \cdot & & \cdot & \cdot \\ 0 & 0 & \cdots & 0 & 1 \end{Vmatrix} = (\alpha_1, \cdots, \alpha_m)\,\bar{\mathrm{A}}.$$

---

[2] T. Levi-Civita, Atti. Ist. Ven. (7), 4 (1892/93) and Rend. Lincei (5), 7 (1898). H. Hahn, Sitz.-Ber. d. Wiener Akademie, Math. Nat. Klasse, Abt. IIa, 116 (1907). A. Ostrowski, *Untersuchungen zur arithmetischen Theorie der Körper*, Math. Zeitschrift, vol. 39 (1934). W. Krull, *Allgemeine Bewertungstheorie*, Crelle vol. 167 (1931). We shall refer to this paper by K.

[3] Cf. K. §5.

Conversely one sees that each matrix

$$\bar{A} = \begin{Vmatrix} 1 & a_{12} & \cdots & \cdots \\ 0 & 1 & a_{23} & \cdots \\ & & 1 & \vdots \\ & 0 & & \ddots & \vdots \\ & & & & 1 \end{Vmatrix}$$

with integers a_{ij} ($i = 1, \cdots, m-1$ and $j = 2, \cdots, m$, $i < j$) defines an automorphism of Γ which conserves the ordering. We note that the determinants of all the representing matrices are equal to one.

Definition. A mapping of the group Γ on a subgroup $\tilde{\Gamma}$ which is isomorphic with Γ and of finite index in Γ is called a *meromorphism* M of Γ.

Such a meromorphism M can again be represented—with respect to a fixed basis of Γ—by a matrix with integral coefficients. Since the ordering must remain unchanged, we have $\Delta_i \cap \tilde{\Gamma} = \tilde{\Delta}_i$ for every isolated subgroup $\tilde{\Delta}_i$ of $\tilde{\Gamma}$. Therefore we obtain the following set of relations

$$\begin{aligned} \tilde{\alpha}_1 &= \alpha_1 b_{11} \\ \tilde{\alpha}_2 &= \alpha_1 b_{12} + \alpha_2 b_{22} \\ &\cdots\cdots\cdots\cdots \\ \tilde{\alpha}_m &= \alpha_1 b_{1m} + \alpha_2 b_{2m} + \cdots + \alpha_{m-1} b_{m-1m} + \alpha_m b_{mm}, \end{aligned}$$

where the coefficients b_{ij} in the diagonal are positive integers. We write $\Gamma \times \bar{M} = (\alpha_1, \cdots, \alpha_m)\bar{M} = (\tilde{\alpha}_1, \cdots, \tilde{\alpha}_m) = \tilde{\Gamma}$. The index of $\tilde{\Gamma}$ in Γ is given by the product $\prod_{i=1}^{m} b_{ii}$ which shall be called the norm $\mathfrak{n}(\bar{M})$ of $\bar{M}$. Obviously all representing matrices $\bar{M}$ of a meromorphism M have the same elements b_{ii} in the diagonal, therefore $\mathfrak{n}(\bar{M})$ is independant of the special representation and an invariant of M.

Conversely each matrix $\bar{M}$ with positive integral coefficients in the diagonal defines a meromorphism with respect to some fixed basis of Γ.

The meromorphism M of Γ on $\tilde{\Gamma}$ can be changed by automorphisms of Γ and $\tilde{\Gamma}$. This means that the representing matrix of M can be transformed into $\bar{A}\bar{M}\bar{\bar{A}}$ where $\bar{A}$ and $\bar{\bar{A}}$ correspond to automorphisms of Γ and $\tilde{\Gamma}$. Therefore the class $<\bar{A}\bar{M}\bar{\bar{A}}>$ of all matrices representing the meromorphism M of Γ upon $\tilde{\Gamma}$ with respect to arbitrary bases is really the equivalent of the abstract mapping $\Gamma \to \tilde{\Gamma}$.

One easily verifies that to the norm of the product of two abstract meromorphisms there corresponds the product of the norms.

The mapping of Γ on the subgroup $n \cdot \Gamma$ of all elements of Γ multiplied by an integer n has the class $<n \cdot \bar{A}>$. The index $[\Gamma : n \cdot \Gamma]$ is equal to n^m.

2. Maximally Perfect Fields and Fields of Formal Power Series in Several Variables

First we wish to state some well known facts on maximally perfect fields.[4] Let K be a maximally perfect field with respect to a valuation $\mathbf{V}$ of rank m; let $\Gamma = \Gamma(K) = \Delta_m$ denote the associated valuation group. All elements a in K whose values $\mathbf{v}(a)$ are non-negative form a ring $R[\mathbf{V}] = R[\Gamma] = R[\mathfrak{p}(0)]$ in K, the so-called $\mathbf{V}$-ring of K; the subset $\mathfrak{p}(0)$ of all elements $b \in R[\mathbf{V}]$ whose values are greater than 0 form a prime ideal. We now associate to each element of $R[\mathbf{V}]$ its residue class modulo $\mathfrak{p}(0)$, $R[\mathbf{V}]/\mathfrak{p}(0) = K(0) \cong K'$ is a field. This mapping is a homomorphism of a subring of K upon K'. Furthermore we let correspond to all elements $c \in K$ which do not lie in $R[\mathbf{V}]$ a formal symbol ∞; obviously we have $\mathbf{v}(c) < 0$ for all these elements. We then say that K is mapped homomorphically upon $\{K', \infty\}$.[5]

To the different isolated subgroups Δ_{m-i} of Γ there correspond also homomorphisms of $K = K(m)$ upon fields $K(m - i)$ and symbols ∞. These mappings are obtained as follows. Take a fixed subgroup Δ_{m-i}. Then all elements $d \in R[\mathbf{V}]$ whose values do not lie in Δ_{m-i} form a prime ideal $\mathfrak{p}(m - i)$ in $R[\mathbf{V}]$. We have $\mathfrak{p}(m - 1) \subset \mathfrak{p}(m - 2) \subset \cdots \subset \mathfrak{p}(1) \subset \mathfrak{p}(0)$. The elements a' of the form a/d' with $a \in R[\mathbf{V}]$ and $d' \neq 0$ prime to $\mathfrak{p}(m - i)$ form a ring $R(\mathfrak{p}(m - i)) = R[\Gamma/\Delta_{m-i}] = R(m - i)$ containing $R[\mathbf{V}]$. Then $\mathfrak{p}(m - i)$ is also a prime ideal in $R(\mathfrak{p}(m - i))$. This ring is a valuation ring with the value group Γ/Δ_{m-i}; let the valuation which belongs to it be called $\mathbf{V}(i)$, i is the rank of Γ/Δ_{m-i}. As above the process of forming the residue classes of $R(m - i)$ modulo $\mathfrak{p}(m - i)$ defines a homomorphism of $K = K(m)$ upon $\{R(m - i)/\mathfrak{p}(m - i) = K(m - i), \infty\}$. The isolated subgroup Δ_{m-i} belongs to a valuation of rank $m - i$ in $K(m - i)$; its valuation ring is equal to $R[\mathbf{V}]/\mathfrak{p}(m - i)$. The field $K(m - i)$ is mapped upon $\{K = K', \infty\}$ by this valuation. Furthermore $K(m - i)$ is mapped upon $\{K(m - i - 1), \infty\}$ by the archimedean valuation corresponding to the factor group $\Delta_{m-i} \mid \Delta_{m-i-1}$ which is isomorphic to a subgroup of the additive group of all real numbers. If Γ is discrete then the latter are isomorphic with the additive group of all integers.

The field K is also perfect with respect to the valuation $\mathbf{V}(1)$ of rank one. This can be shown as follows. Assume that K could be imbedded in a larger field $\tilde{K}$ which is perfect with respect to $\mathbf{V}(1)$ and which has the same residue class field $K(m - 1)$. Then $\tilde{K}$ would be perfect with respect to a valuation $\tilde{\mathbf{V}}$ of rank m and the residue class field $\tilde{K}'$ would be equal to K'. According to the decomposition of value groups the associated group $\tilde{\Gamma}$ would be equal to Γ. That means that $\tilde{K}$ is an immediate extension of K with respect to the valuation $\mathbf{V}$. But this is impossible because K was assumed to be maximally perfect.

---

[4] Cf. K. §§1, 5.

[5] Conversely each homomorphism φ of K on a field K' and a symbol ∞ determines uniquely a valuation $\mathbf{V}$ of K, if $a\varphi = \infty \leftrightarrow a^{-1}\varphi = 0$. All elements $r \in K$ with $r\varphi \in K'$ form a valuation ring R in K, and all elements s with $s\varphi = 0$ form a prime ideal in R. It is quite obvious how to construct the valuation and its group. See e.g. K. §1.

THEOREM 1. *A maximally perfect field K with respect to a valuation $\mathbf{V}$ of rank m which is not algebraically closed, is maximally perfect with respect to the valuation $\mathbf{V}$ only.*

PROOF. First we show that distinct valuations $\mathbf{V}^1$ and $\mathbf{V}^2$ of rank m belonging to a field K generate distinct valuations $\mathbf{V}(1)^1$ and $\mathbf{V}(1)^2$ of rank one. We assume that $\mathbf{V}(1)^1$ and $\mathbf{V}(1)^2$ are equivalent. Then there exist always suitable automorphisms of the groups Γ^1 and Γ^2 belonging to $\mathbf{V}^1$ and $\mathbf{V}^2$ such that any element of K has the same values in the respective groups Γ^1/Δ^1_{m-1} and Γ^2/Δ^2_{m-2} which belong to the equivalent valuations $\mathbf{V}(1)^1$ and $\mathbf{V}(1)^2$. According to a fundamental property of valuations there exist elements a and b in K such that their values with respect to $\mathbf{V}^1$ have zero components with respect to the subgroup Δ^1_{m-1} and only components in Γ^1/Δ^1_{m-1}.[6] These elements can furthermore be chosen in such a way that $\mathbf{v}^1(a) > \mathbf{v}^1(b)$ and that $\mathbf{v}^2(a) < \mathbf{v}^2(b)$. Then we have also $\mathbf{v}(1)^1(a) > \mathbf{v}(1)^1(b)$ according to the special choice of a and b. According to the assumption on $\mathbf{V}(1)^1$ and $\mathbf{V}(1)^2$ and on the normalization of the groups Γ^1 and Γ^2 we would have $\mathbf{v}(1)^2(a) > \mathbf{v}(1)^2(b)$. Hence according to the synthesis of the valuation groups Γ^1 and Γ^2 out of Γ^1/Δ^1_{m-1}, Δ^1_{m-1} and Γ^2/Δ^2_{m-1}, Δ^2_{m-1} we would obtain $\mathbf{v}^2(a) > \mathbf{v}^2(b)$, in contradiction to the distinctness of $\mathbf{V}^1$ and $\mathbf{V}^2$.

We next assume that K is maximally perfect with respect to two inequivalent valuations $\mathbf{V}^1$ and $\mathbf{V}^2$ of rank m. Then K is also perfect with respect to the two distinct valuations $\mathbf{V}(1)^1$ and $\mathbf{V}(1)^2$ of rank one. According to a theorem of F. K. Schmidt this can only happen if K is algebraically closed and if a certain relation for the power of K holds.[7] The latter condition can be discarded in our case. The first condition is certainly not fulfilled because K was assumed to be not closed.

From now on we suppose that the field K under consideration has the *same characteristic as its field of residue classes K'* which belongs to a valuation of K. Let $\hat{K} = K(K', \Gamma) = \{\sum_{\gamma_{i+1} > \gamma_i > \cdots > \gamma_0} k'_i u^{\gamma_i}\}$ be the field of all symbolic power series with coefficients in the field K' and exponents γ_i in a discrete ordered group Γ of rank m. The field $\hat{K}$ contains the ring $\hat{R}[\Gamma]$ of all power series of the form $k'_0 + \sum_{\gamma_i > 0} k'_i u^{\gamma_i}$. The power series with vanishing k'_0 form a prime ideal $\hat{\mathfrak{p}}$ in $\hat{R}[\Gamma]$ because the ring $\hat{R}[\Gamma]/\hat{\mathfrak{p}} \cong K'$ contains no zero divisors. Each element $a \in \hat{R}[\Gamma]$ can be written in the form $a = u^{\alpha}(k'_0 + \sum_{\gamma_i > 0} k'_i u^{\gamma_i})$ with $k'_0 \neq 0$. We then define a value function $\mathbf{v}$ on $\hat{K}$ by $\mathbf{v}(a) = \alpha$. Obviously $\mathbf{v}$ fulfills all postulates of a value function belonging to a valuation $\mathbf{V}$ with the group Γ. For arbitrary elements $b \in \hat{K}$ which do not lie in $\hat{R}(\Gamma)$ there exists a power u^{β} such that $bu^{\beta} = k'_0 + \sum_{\gamma_i > 0} k'_i u^{\gamma_i}$ with $k'_0 \neq 0$; we then put $\mathbf{v}(b) = -\beta$.

Now we define abstract symbols t_i $(i = 1, 2, \cdots, m)$ to which we attribute the values α_i. The product $\prod_{i=1}^{m} t_i^{a_i}$ shall have the value $\sum_{i=1}^{m} a_i\alpha_i$, and furthermore we postulate that there exist no additive relations between these products.

---

[6] Cf. K. §10, Th. 15.

[7] F. K. Schmidt, *Mehrfach bewertete Körper*, Math. Annalen vol. 108 (1933).

Then we can map the field $\hat{K}$ on the set of all power series in the m variables t_i by the rule

$$\sum_{\gamma_i=\Sigma a_{ij}\alpha_j>0} k_i' u^{\gamma_i} \to \sum k_i' t_1^{a_{i1}} \cdots t_m^{a_{im}}.$$

According to the construction of the t_i these power series form a field K which is isomorphic with the field $\hat{K}$ of all symbolic power series. We have then proved the following lemma.

LEMMA 1. *The field of all power series* $K = \{\sum k_i' t_1^{a_{i1}} \cdots t_m^{a_{im}}\}$ *where the coefficients* k_i' *are taken from a field* K', *and where the systems of exponents run over the coefficients of an ascending sequence of elements belonging to a discrete ordered abelian group* Γ *of rank* m, *is a maximally perfect field with the valuation group* Γ *and the field* K' *as field of residue classes.*

Now let K be a maximally perfect field with respect to a discrete valuation $\mathbf{V}$ of rank m, and let the field of residue classes belonging to K with respect to $\mathbf{V}$ be K'. By employing the m isolated subgroups of Γ we can decompose the homomorphism $K \to \{K', \infty\}$ in m homomorphisms $K(m - i) \to \{K(m - i - 1), \infty\}$ of rank one. The fields $K(m - i)$ are perfect with respect to discrete valuations; $K(m - i - 1)$ are the corresponding fields of residue classes and $K(m - i)$ are not algebraically closed because Δ_{m-i} are discrete, $(i \neq m)$. According to the structure of perfect fields with respect to archimedean discrete valuations the field $K = K(m)$ is isomorphic with a field $\hat{K} = \hat{K}(m)$ of power series in one variable t_m with coefficients in $K(m - 1)$, and $\hat{K}$ contains a field $\hat{K}(m - 1)$ which is isomorphic with $K(m - 1)$. The field $K(m - 1)$ is perfect with respect to a discrete archimedean valuation and has as field of residue classes the field $K(m - 2)$. Again $K(m - 1)$ is isomorphic with a field $\hat{K}(m - 1)$ of power series in one variable t_{m-1} with coefficients in $K(m - 2)$, and $K(m - 2)$ can be imbedded isomorphically as $\hat{K}(m - 2)$ in $\hat{K}(m - 1)$. This process can be continued to $K(1) \to \{K(0) = K', \infty\}$. Finally we obtain a field $\hat{K}$ which is isomorphic with the given field K. The field $\hat{K}$ contains isomorphic images of the homomorphic residue class fields $K(m - i)$, and $K(m - i)$ is always isomorphic with a field of power series in one variable with coefficients in an imbedded subfield which is isomorphic with $K(m - i - 1)$. One easily sees by developing the coefficients successively that the field $\hat{K}$ is isomorphic with a field of power series in m variables with coefficients in a subfield which is isomorphic with K'. Furthermore the systems of exponents are ordered in the sense of the valuation group Γ belonging to $\mathbf{V}$. Hence we have the following theorem.

THEOREM 2. *Each maximally perfect field* K *with respect to a discrete valuation* $\mathbf{V}$ *of rank* m *and the field* K' *of residue classes is isomorphic with a field of power series* $\hat{K}$ *in* m *variables with coefficients in* K'.[8]

COROLLARY. *Each maximally perfect field* K *with respect to a discrete valua-*

[8] The field of coefficients is in general not uniquely determined in K.

tion $\mathbf{V}$ *of rank* m *is defined uniquely up to abstract isomorphisms by the field of residue classes* K' *and the valuation* $\mathbf{V}$.

This is a consequence of Theorem 1, Lemma 1 and Theorem 2.

3. Arithmetic and Algebraic Properties of Maximally Perfect Fields

Let K be a perfect field with respect to the discrete valuation $\mathbf{V}$ of rank m, and let the associated group of values be called $\Gamma(K)$. By L we shall denote some finite algebraic extension of degree n over K. Then it is a well known fact that L is also perfect with respect to a discrete valuation $\mathbf{V}$ of rank m whose value group $\Gamma(L)$ contains $\Gamma(K)$. The value of an element a in L is defined by the n^{th} part of the value belonging to its norm $\mathfrak{N}(L, K \mid a)$.[9] Therefore the index $[\Gamma(L):\Gamma(K)]$ is finite and a divisor of n^m. The group $\Gamma(K)$ is then a meromorphic map of $\Gamma(L)$, because both groups are abstractly isomorphic.

Definition. The class $cl(\bar{\mathrm{E}})$ of all meromorphisms mapping $\Gamma(L)$ into $\Gamma(K)$ shall be called the *ramification class* of L over K.

The values of the norms $\mathfrak{N}(L, K \mid a)$ of all elements in L form a subgroup of $\Gamma(K)$. They are equal to $n \cdot \mathbf{v}_L(a)$ because conjugate elements have the same values. (Eventually we have to imbed L in its smallest normal extension over K, but that can be done without loss of generality as one easily observes.) We may write without danger of confusion for the group of values belonging to the norms $\mathbf{v}_K(\mathfrak{N}(L, K \mid a \in L)) = n \cdot \mathbf{v}_L(a \in L) = \mathfrak{N}(L, K \mid \Gamma(L))$. This group has a finite index under $\Gamma(K)$ and is abstractly isomorphic with $\Gamma(K)$, therefore it is a meromorphic map of $\Gamma(K)$.

Definition. The class $cl(\bar{\Phi})$ of all meromorphisms mapping $\Gamma(K)$ into $N(L, K \mid \Gamma(L))$ shall be called the *norm class* of L over K.

Lemma 2. *For any finite algebraic extension L of degree n with respect to K there holds the relation* $cl(\bar{\mathrm{E}})cl(\bar{\Phi}) = cl(n \cdot \bar{\mathrm{A}})$.

Proof. According to the definitions of $cl(\bar{\mathrm{E}})$, $cl(\bar{\Phi})$ and the inclusion $\mathfrak{N}(L, K \mid \Gamma(L)) \subseteq \Gamma(K) \subseteq \Gamma(L)$ the assertion of the theorem becomes evident. For $\mathfrak{N}(L, K) \mid \Gamma(L) = \Gamma(K) \times \bar{\Phi} = \Gamma(L) \times \bar{\mathrm{E}} \times \bar{\mathrm{A}} \times \bar{\Phi} = n \cdot \Gamma(L)$. We abbreviate that relation symbolically by $cl(\bar{\mathrm{E}})cl(\bar{\Phi}) = cl(n \cdot \bar{\mathrm{A}})$. We understand this equation to mean that for given $\bar{\mathrm{E}}$ and Φ in the classes $cl(\bar{\mathrm{E}})$ and $cl(\bar{\Phi})$ there exist representing matrices of automorphisms $\bar{\mathrm{A}}_1$, $\bar{\mathrm{A}}_2$, $\bar{\mathrm{A}}_3$, $\bar{\mathrm{A}}_4$ and $\bar{\mathrm{A}}_5$ such that $\bar{\mathrm{A}}_1\bar{\mathrm{E}}\bar{\mathrm{A}}_2\bar{\mathrm{A}}_3\bar{\Phi}\bar{\mathrm{A}}_4 = n \cdot \bar{\mathrm{A}}_5$.

The elements in the diagonals of all meromorphisms belonging to the same class are obviously invariants. For the norms $\mathfrak{n}(cl(\bar{\mathrm{E}}))$ and $\mathfrak{n}(cl(\bar{\Phi}))$ of the ramification and norm class belonging to L over K we obtain the equation $\mathfrak{n}(cl(\bar{\mathrm{E}}))\mathfrak{n}(cl(\bar{\Phi})) = n^m$. This relation will be of special importance later on.

Remark. For $m = 1$ the characteristic classes $cl(\bar{\mathrm{E}})$ and $cl(\bar{\Phi})$ of a field L over K become positive integers e and f. They are the same as the ramifica-

[9] We introduce $N(L, K \mid a)$ for the norm of an element $a \in L$ taken from L relative to K in order to simplify the printing.

tion and residue class degrees as introduced in the theories of algebraic numbers and algebraic functions of one variable.

It is possible to introduce the notions of ramification number and residue class degree for fields which are perfect with respect to a valuation of rank m (> 1). For this purpose we consider the valuation $\mathbf{V}(1)$ of K which belongs to the factor group Γ/Δ_{m-1}. The associated prime ideal $\mathfrak{p}(m-1)$ is generated by any element of K whose value is equal to $\alpha_1 a_1 + \cdots + \alpha_{m-1} a_{m-1} + \alpha_m$. Then we have $\mathfrak{p}(m-1) = \mathfrak{P}^e(m-1)$ and $R[L, \mathfrak{P}(m-1)]/\mathfrak{P}(m-1)$: $R[K, \mathfrak{p}(m-1)/\mathfrak{p}(m-1)] = f$, where $ef = n$.[10] Incidentally $e = e_m$ as one easily sees from the decomposition of the valuation $\mathbf{V}$ into $\mathbf{V}(1)$ and a valuation of rank $m-1$ in $K(m-1)$. Later we shall also give interpretations of the remaining coefficients $e_1, \cdots, e_m$ of the ramification class, but there it will be necessary to make restricting assumptions on the field K' of residue classes.

Let us assume for the moment that L is a cyclic extension of K whose degree n is prime to the characteristic of K, and that K contains the n^{th} roots of unity. Then L is equal to the radical field $K(a^{1/n})$ for a suitably chosen element a in K. In this case the ramification number $e = e_m$ can be determined by the value of a.

Lemma 3. *The ramification number of a cyclic field $L = K(a^{1/n})$ of degree n is given by the index $[\{a_m/n, \Gamma^0\}:\Gamma^0]$ where $\mathbf{v}(a)$ is equal to $\alpha_1 a_1 + \cdots + \alpha_m a_m$.*

Proof. We reduce the group Γ of K modulo Δ_{m-1}. For the archimedean valuation $\mathbf{V}(1)$ belonging to the factor group Γ/Δ_{m-1} the calculation is easily carried out.[11]

In all the following investigations we restrict ourselves to maximally perfect fields *K whose residue class fields are algebraically closed fields of characteristic zero.*

Theorem 3. *Every finite algebraic extension L over K is abelian, and the Galois group is isomorphic with the factor group $\Gamma(L)/\Gamma(K)$.*

Proof. It suffices to show that every normal extension L of K has an abelian Galois group, because each field can be imbedded in a smallest normal field; and the fact that all normal fields are abelian implies according to the Galois theory that all subfields are also abelian. To prove this fact we make use of the generalization of Hilbert's ramification theory as developed by Krull.[12]

First we observe that K coincides with the decomposition field belonging to the valuation $\mathbf{V}$ and any normal extension L. This is a consequence of the perfectness of K. Second K is equal to the field of inertia belonging to $\mathbf{V}$ with respect to L. The group associated with such an inertial field is also the Galois group of the residue class field $R[\Gamma(L)]/\mathfrak{P}(0)$ with respect to $R[\Gamma(K)]/\mathfrak{p}(0)$. Both fields are isomorphic with the algebraically closed field K'. Therefore there cannot exist a proper group of inertia. Then L is exactly the field of

---

[10] The relation $ef = n$ holds always for discrete archimedean valuations. See for ex. M. Deuring, *Algebren* (Berlin 1935). Ch. VI, §12.

[11] Cf. M. Deuring, *Verzweigungstheorie bewerteter Körper*, Math. Annalen vol. 105 (1931).

[12] See K. §9.

ramification over K with respect to the valuation $\mathbf{V}$. Fields of higher ramification do not exist because K and K' are of characteristic zero. Hence according to a general theorem of Krull the field L is abelian over K and its Galois group is isomorphic with $\Gamma(L)/\Gamma(K)$.

REMARK. If $\bar{\mathrm{E}}$ is an arbitrary representation of the ramification class $cl(\bar{\mathrm{E}})$ of L over K, then we have $[L:K] = n = \mathfrak{n}(\bar{\mathrm{E}}) = \mathfrak{n}(cl(\bar{\mathrm{E}}))$. For the norm class we obtain $\mathfrak{n}(cl(\Phi)) = n^{m-1}$.

THEOREM 4. *All extensions L in which the prime ideal $\mathfrak{p}(m-1)$ of K is totally ramified, $\mathfrak{p}(m-1) = \mathfrak{P}^n(m-1)$, are cyclic.*

PROOF. According to theorem 3 the field L is an abelian extension of K. Its Galois group may be represented as the direct product of abelian groups with maximal orders $p^{\varepsilon_p} = \pi$ where π are the exact powers of the primes p dividing the order n of the Galois group. To such a decomposition of the group corresponds a representation of L as a composite of abelian fields L_p of degree π over K. If it can be shown that all the fields L_p are cyclic with respect to K, then their composite L must also be cyclic because their different degrees are relatively prime.

An arbitrary abelian field L_p is a composite of cyclic fields $L_{p,1}, \cdots, L_{p,k}$. Their degrees are proper divisors of π if $r \geqq 2$, that is to say, if L_p is not cyclic. Let the prime ideal $\mathfrak{p}(m-1)$ have the decompositions $\mathfrak{p}(m-1) = \mathfrak{P}^{g(i)}(m-1)$ in these r fields. The numbers $g(i)$ are powers of p and surely proper divisors of π if L_p is not cyclic. Let $\mathfrak{p}(m-1)$ be the $e_p{}^{\text{th}}$ power of the prime ideal $\mathfrak{P}(m-1)$ in L; then $e = n = [L:K]$ is the least common multiple of all the e_p.[13] Hence $e_p = \pi$ for each prime p which divides n. Therefore the fields L_p are also totally ramified. Their ramification numbers e_p are again the least common multiples of the numbers $g(i)$. The latter are proper divisors of $\pi = e_p$ if L_p is not cyclic, hence their least common multiple is also a proper divisor of π. But this leads to a contradiction if $r \geqq 2$. Therefore all the components L_p of a totally ramified field are cyclic.

We are now in a position to give the interpretation previously announced of the coefficients $e_1, \cdots, e_{m-1}, e_m$ which belong to the ramification class of the field L over K.

The homomorphism $K = K(m) \to \{K(0) = K', \infty\}$ may be split off into m homomorphisms $K(m-i) \to \{K(m-i-1), \infty\}$ of rank one. The field $K(m-i)$ is maximally perfect with respect to a valuation of rank $m-i$, the associated value group is Δ_{m-i} and K' is the corresponding field of residue classes. Similar decompositions hold for the field L. In each case $L(m-i) \supseteq K(m-i)$ and $\Delta(L)_{m-i} \supseteq \Delta(K)_{m-i}$.

The prime ideal $\mathfrak{p}(m-1)$ of K becomes an $e_m{}^{\text{th}}$ power of a prime ideal $\mathfrak{P}(m-1)$ in L, as we have seen before. Therefore $[L(m-1):K(m-1)] = ne_m^{-1}$. There exists a meromorphism E_{m-1} mapping $\Delta(L)_{m-1}$ into the whole of

[13] Cf. no. 11.

$\Delta(K)_{m-1}$. According to the structure of ordered discrete groups there exists a representing matrix $\bar{\mathrm{E}}_m$ of the ramification class of L over K such that

$$\bar{\mathrm{E}}_m = \left\| \begin{matrix} & & * \\ & & * \\ \bar{\mathrm{E}}_{m-1} & & \vdots \\ & & * \\ 0 & \cdots\ 0 & e_m \end{matrix} \right\|.$$

Consequently any meromorphism of the ramification class of $L(m-1)$ over $K(m-1)$ has $e_1, \cdots, e_{m-1}$ as elements in the diagonal. The homomorphism $K(m-1) \to \{K(m-2), \infty\}$ determines a prime ideal in $K(m-1)$ which is principal. It is the power of a prime ideal in $L(m-1)$, and the exponent is equal to e_{m-1}. Thus the coefficient e_{m-1} can be interpreted as a ramification number. The degree of the corresponding residue class fields. $[L(m-2):K(m-2)]$ is equal to $[L(m-1):K(m-1)]e_{m-1}^{-1} = \mathfrak{n}(cl(\mathrm{E}_{m-1}))e_{m-1}^{-1} = n(e_m e_{m-1})^{-1}$. The same reasoning can be repeated for the fields $L(m-2)$, $K(m-2)$, etc. Hence we finally see that all the coefficients e_{m-i} are ramification numbers of fields $L(m-i)$ that are homomorphic with L.

There exists still another interpretation of the numbers e_i: they appear as relative degrees of fields in L. These fields are the fields of inertia with respect to the partial valuations pertaining to the factor groups Γ/Δ_{m-i}. Let the field of inertia with respect to the prime ideal $\mathfrak{p}(m-1)$ be called $T(m)_{m-1}$. Its relative degree with respect to K is equal to ne_m^{-1} by what we have seen before; and L is cyclic and totally ramified of degree e_m over $T(m)_{m-1}$. The field $T(m)_{m-1}$ is generated by any equation

$$f(y) = y^{n/e_m} + a_1^{(m-1)} y^{n/e_m - 1} + \cdots + a_{n/e_m}^{(m-1)} = 0$$

where the coefficients $a_j^{(m-1)}$ are representatives of residue classes, $\bar{a}_j^{(m-1)}$. The elements $\bar{a}_j^{(m)} \in K(m-1)$ are the coefficients of any generating equation $y^{n/e_m} + \bar{a}_1^{(m-1)} y^{n/e_m - 1} + \cdots + \bar{a}_{n/e_m}^{(m-1)} = 0$ of the field $L(m-1)$ over $K(m-1)$. For the proof of these facts we observe that $f(y) = 0$ is irreducible in K and that $f(y) = 0$ is totally decomposed in L; this is a consequence of the perfectness of K and L with respect to the partial valuations of rank one. (Application of Hensel's criterion of reducibility.)[14] One also sees easily that $T(m)_{m-1}$ is the largest unramified field in L such that its field of residue classes with respect to $P(m-1)$ is isomorphic with $L(m-1)$.

By the same methods it is shown that $L(m-1)$ contains a maximal unramified field $T(m-1)_{m-2}$ with respect to the prime ideal belonging to the homomorphism $L(m-1) \to \{L(m-2), \infty\}$. For the relative degree $[T(m-1)_{m-2}:K(m-1)]$ we obtain $n(e_m e_{m-1})^{-1}$, and $L(m-1)$ is totally ramified of degree e_{m-1} over $T(m-1)_{m-2}$. According to Hensel's criterion there exists a uniquely determined subfield $T(m)_{m-2}$ in $T(m)_{m-1}$ such that the

[14] Cf. K. §7. Th. 12.

prime ideal $\mathfrak{p}(m-2)$ is unramified in $T(m)_{m-2}$, that is to say $\mathfrak{p}(m-2)$ of K remains a prime ideal in $T(m)_{m-2}$, and such that the field of residue classes belonging to $T(m)_{m-2}$ modulo $\mathfrak{p}(m-2)$ is equal to $L(m-2)$. The field $T(m)_{m-2}$ is obviously the inertial field belonging to $\mathbf{V}(2)$ in L. The relative degree $[T(m)_{m-1}:T(m)_{m-2}]$ is equal to e_{m-1}.

By iterating this construction we obtain in L a descending series of fields $T(m)_{m-i}$ which are the inertial fields belonging to the valuations $\mathbf{V}(i)$ and the prime ideals $\mathfrak{p}(m-i)$. Their relative degrees $[T(m)_{m-i+1}:T(m)_{m-i-1}]$ are equal to the coefficients e_{m-i} of the ramification class belonging to L over K. Thus these coefficients can be interpreted as degrees of fields in L.

REMARK. The fields $T(m)_{m-i}$ are isomorphic with fields of power series in i variables with coefficients lying in $L(m-i)$. This is an immediate consequence of the preceding developments and the corollary of no. 1. Therefore $L(m-i)$ is isomorphically imbedded in $T(m)_{m-i}$.

The degree $[L:T(m)]$ is also given by the norm of the ramification class of L over K, if K is considered as perfect field with respect to the valuation $\mathbf{V}(i)$. If we denote a corresponding meromorphism of the ramification class by $\mathrm{E}(\Gamma/\Delta_{m-i})$, then $[L:T(m)_{m-i}] = \mathfrak{n}(cl(\mathrm{E}(\Gamma/\Delta_{m-i})))$. The Galois group of L over $T(m)_{m-i}$ is isomorphic with the factor group $\Gamma(L) \mid \Delta(L)_{m-i} \mid \Gamma(K) \mid \Delta(K)_{m-i}$. The group of $T(m)_{m-i}$ over K is isomorphic with $\Delta(L)_{m-i} \mid \Delta(K)_{m-i}$ and $[T(m)_{m-i}:K] = n\mathfrak{n}(\mathrm{E}(\Gamma/\Delta_{m-i}))^{-1}$.

EXAMPLE. If $L = K(t_1^{1/e_1}, \cdots, t_m^{1/e_m})$ then the ramification class of L over K can be represented by the matrix

$$\left\| \begin{matrix} e_1 & & \\ & \ddots & \\ & & e_m \end{matrix} \right\|.$$

The Galois group of L with respect to K is the direct product of m cyclic groups $\mathfrak{Z}(e_i)$ of order e_i. The unramified fields $T(m)_{m-i}$ are given by $K(t_1^{1/e_1}, \cdots, t_i^{1/e_i})$.

4. Further Algebraic Properties of Maximally Perfect Fields

We have already seen that all finite algebraic extensions L over a maximally perfect reference field K with an algebraically closed field K' of residue classes are abelian. Therefore all such fields are composites of radical fields. We shall now make a survey of all superfields L of a fixed degree n over K.

THEOREM 5. *All units of K with respect to the valuation $\mathbf{V}$ of rank m are n-th powers.*

PROOF. For this purpose it is necessary to show that the equation $x^n - \epsilon = 0$, with an arbitrary unit ϵ of K, is reducible. Each unit ϵ is congruent to an element $a' \neq 0$ of the residue class field K'. The equation $x^n - a' = 0$ is reducible in the algebraically closed field K' and its linear factors are different. Repeated application of the generalized criterion of Hensel leads to a decom-

position $x^n - \epsilon = \prod_{j=1}^{n} (x - \epsilon_j)$, where the units e_j are products of an arbitrary ϵ_1 amongst them with suitable n^{th} roots of unity.

The n^{th} roots of any two elements a_1 and a_2 differing only by a unit ϵ with respect to the valuation $\mathbf{V}$ generate one and the same radical field $L = K(a_1^{1/n}) = K(a_2^{1/n})$. All elements $a \neq 0$ of K whose values $\mathbf{v}(a)$ are equal to

$$\alpha_1 b_1 \cdot n + \cdots + \alpha_m b_m \cdot n$$

are n^{th} powers of elements b in K^*, if we denote by K^* the multiplicative group of K. Therefore all different radical fields of degree n are generated by elements a whose values $\mathbf{v}(a)$ are representatives of the residue classes of $\Gamma/n\cdot\Gamma$. We have proved the following theorem.

THEOREM 6. *The factor group K^*/K^{*n} is isomorphic with the factor group $\Gamma/n\cdot\Gamma$ which is the direct product of m cyclic groups of order n.*

Previously we have shown that the Galois group of a field L over K is isomorphic with the factor group $\Gamma(L)/\Gamma(K)$ of the respective value groups. The group $\Gamma(L)$ can be considered as an (ordered) extension of the group $\Gamma(K)$ contained in the group $1/n$, $\Gamma(K)$ consisting of all linear forms

$$1/n\cdot a_1\alpha_1 + \cdots + 1/n\cdot a_m\alpha_m .$$

The latter is the value group belonging to the field $K(t_1^{1/n}, \cdots, t_m^{1/n})$ which contains the given field L of degree n according to the theory of radical fields.

THEOREM 7. *The extensions L of degree n over a maximally perfect field K correspond uniquely to the different ordered subgroups of order n with respect to $\Gamma(K)$ contained in the group $1/n\cdot\Gamma(K)$.*

PROOF. Let $\tilde{\Gamma}$ be an ordered extension of order n over $\Gamma(K)$. It can be generated by m linear forms $\tilde{\alpha}_i = \alpha_1 r_{i1} + \cdots + \alpha_m r_{im}$ where the numbers r_{ij} are rational fractions which have at most the denominator n. This follows from the fact that $\tilde{\Gamma}$ is contained in $1/n\cdot\Gamma(K)$. Furthermore the matrix $\| r_{ij} \|$ can be transformed into the diagonal form with positive elements r_{ii} by a suitably chosen automorphism of $\Gamma(K)$ because Γ was assumed to be an ordered group. Now we adjoin the elements $T_i = t_1^{r_{i1}} \cdots t_m^{r_{im}}$, with respect to a fixed normalization of the n^{th} root, to the field K, thus obtaining a field $\tilde{\Gamma}$ of degree n over K. The values of the elements T_i with respect to the valuation $\tilde{\mathbf{V}}$ of $\tilde{L}$ are equal to $1/n\cdot\mathbf{v}(\mathfrak{N}(L, K \mid T_i)) = \tilde{\alpha}_i$ by definition of the valuation $\tilde{\mathbf{V}}$ in the algebraic extension of K. According to theorem 3 $\Gamma(\tilde{L})/\Gamma(K)$ is of order n, therefore the elements $\tilde{\mathbf{v}}(T_i)$ which form an extension of order n over $\Gamma(K)$ already generate the whole group $\Gamma(\tilde{L})$. The inverse of a matrix representing the ramification class $cl(\mathrm{E})$ of $\tilde{L}$ over K is obviously equivalent to $\| r_{ij} \|^{-1}$ with respect to matrices representing automorphisms of the value groups. Conversely, the system $\| r_{ij} \|^{-1}$ belonging to a given field L of degree n over K belongs to the class $cl(\mathrm{E})$. We have $\{\| r_{ij} \|\} = \{n^{-1}\cdot\bar{\Phi}\}$ according to lemma 2.

According to the theory of radical fields it is evident that different extensions $\tilde{\Gamma}_1$ and $\tilde{\Gamma}_2$ contained as $1/n\cdot\Gamma(K)$ generate different fields. Thus we see that the ramification class uniquely determines the field to which it belongs.

REMARK. A remarkable consequence of these observations (combined with theorem 6) is that there exists only a finite number of types of ramification classes for all extensions of a fixed degree n over fields F of algebraic functions of m variables. We recall that the different valuations $\mathbf{V}_1, \cdots, \mathbf{V}_g$ of a finite extension M of F are given by the simple components M_i of the direct product $M \times F_{\mathbf{V}}$ of the field M with the maximally perfect extension $F_{\mathbf{V}}$ of F with respect to the valuation $\mathbf{V}$ of rank m.[15] It follows that $g \leqq n$ and $[M_i:F_{\mathbf{V}}] \leqq n$.

5. Cyclic Fields and Groups of Norms

In this section we shall investigate specially the cyclic extensions of maximally perfect fields K with respect to a discrete valuation of rank m. We shall see that there exist certain generalizations of the local class field theory.

We assume that all fields occurring in the following considerations are contained in one and the same algebraic, algebraically closed extension of the reference field K.

The group of all elements unequal to 0 in K which are norms of elements in L shall be denoted by $\mathfrak{N}(L, K \mid L^*)$; if there is no danger of ambiguity we merely write $\mathfrak{N}(L^*)$.

THEOREM 8. *The order of the norm class group $K^*/\mathfrak{N}(L^*)$ of any extension L of degree n over K is equal to n^{m-1}.*

PROOF. According to theorem 5 all units of K are n^{th} powers of units. Therefore they are also norms of units in L. Each element $a \in K^*$ is equal to $a_0\epsilon$, where $\mathbf{v}(a) = \mathbf{v}(a_0)$ and $a_0 = \prod_{i=1}^m t_i^{a_i}$. The elements t_i form a system of elements whose values are equal to α_i. Thus we obtain the isomorphisms $K^*/N(L^*) \cong \Gamma(K)/\mathfrak{N}(L, K \mid \Gamma(L)) = \Gamma(K)/\Gamma(K) \times \Phi$. Hence $[K^*:\mathfrak{N}(L^*)] = [\Gamma(K):\Gamma(K) \times \Phi] = \mathfrak{n}(cl(\Phi))$. According to lemma 2 and theorem 3 we have $\mathfrak{n}(cl(\Phi)) = n^m \mathfrak{n}(cl(\mathrm{E}))^{-1} = n^{m-1}$.

REMARK. For any algebraic extension L of degree n over K we obtain $[\mathfrak{N}(L^*):K^*] = n$.

In the case that $L = Z$ is a cyclic extension of K we are able to describe the structure of the group $K^*/\mathfrak{N}(Z^*)$.

THEOREM 9. *The norm class group $K^*/\mathfrak{N}(Z^*)$ of a cyclic superfield Z of degree n over K is the direct product of $m - 1$ cyclic groups of order n.*

PROOF. We first determine the structure of the norm class group for cyclic fields $Z(p)$ whose degree are powers of a single prime p, $[Z(p):K] = p^{\nu_p} = \pi$.

The field $Z(p)$ is generated by the radical $(t_1^{d_1} \cdots t_i^{d_i} \cdots t_m^{d_m})^{1/\pi}$. The greatest common divisor d of the exponents $d_1, \cdots, d_i, \cdots, d_m$ is prime to p if $Z(p)$ has exactly the degree π over K. Put $d^{-1} = \delta$. According to the theory of radical fields the cyclic field $Z(p)$ is also generated by the radical

$$(t_1^{d_1\delta} \cdots t_i^{d_i\delta} \cdots t_m^{d_m\delta})^{1/\pi}.$$

[15] In order to determine a valuation of F by a prime ideal $p(m - i)$ of a maximal order $\mathfrak{o}$ of F it is necessary to consider a chain of prime ideals which are multiples of $\mathfrak{p}(m - i)$ in $\mathfrak{o}$, that is to say a chain $0 \subset \mathfrak{p}(m-1) \subset \mathfrak{p}(m-2) \subset \cdots \subset \mathfrak{p}(m-i)$. Only so is it possible to construct an ordering in the quotient ring belonging to $\mathfrak{p}(m - i)$ in F.

We have $(d_1 d^{-1}, \cdots, d_i d^{-1}, \cdots d_m d^{-1}) = 1$. As a consequence of $[Z(p):K] = \pi$ there exists at least one exponent $d_i d^{-1} = \delta(i)$ which is prime to p. The additive congruence $\delta(i) \cdot x \equiv 1 \pmod{\pi}$ has therefore an integral solution x which is prime to p. Now we form the radical $(t_1^{\delta(1)x} \cdots t_i^{\delta(i)x} \cdots t_m^{\delta(m)x})^{1/\pi}$; it also generates the field $Z(p)$. The same holds for the radical $(t_1^{\delta(1)x} \cdots t_i^1 \cdots t_m^{\delta(m)x})^{1/\pi}$ because $\delta(i)x \equiv 1 \pmod{\pi}$. The group $K^* = \{t_1, \cdots, t_i, \cdots, t_m\} \times \epsilon(K)$, $\epsilon(K)$ standing symbolically for the group of all units of K, can be represented by the equivalent basis $\{t_1, \cdots, t_{i-1}, t_1^{\delta(1)x} \cdots t_i^1 \cdots t_m^{\delta(m)x}, t_{i+1}, \cdots, t_m\} \times \epsilon(K)$.

Now $t_1^{\delta(1)x} \cdots t_i^1 \cdots t_m^{\delta(m)x} = \mathfrak{N}(Z(p), K \mid (t_1^{\delta(1)x} \cdots t_i^1 \cdots t_m^{\delta(m)x})^{1/\pi})$. The group of all norms $\mathfrak{N}(Z(p)^*)$ therefore contains the subgroup

$$S = \{t_1^\pi, \cdots, t_1^{\delta(1)x} \cdots t_i^1 \cdots t_m^{\delta(m)x}, \cdots, t_m^\pi\} \times \epsilon(K).$$

That group has the index π^{m-1} in K^*, and the factor group K^*/S is of type $(\pi, \cdots, \pi)$ with $m - 1$ terms π. According to our construction $S \subseteq \mathfrak{N}(Z(p)^*)$ and consequently $S = \mathfrak{N}(Z(p)^*)$. Hence

$$K^*/\mathfrak{N}(Z(p)^*) \cong \mathfrak{Z}(\pi)_1 \times \cdots \times \mathfrak{Z}(\pi)_{m-1},$$

with isomorphic cyclic groups $\mathfrak{Z}(\pi)_i$ of order $\pi = p^{\nu_p}$.

Now let be $n = p^{\nu_p} \cdot r^{\nu_r} \cdots q^{\nu_q} = \pi \cdot \rho \cdots \kappa$. In order to simplify the notations we treat only the case $n = \pi\rho$; one easily sees by considering the following proof that the general case can be treated by induction with respect to the different divisors of n.

The cyclic field $Z = K(a^{1/n})$ contains the cyclic fields $Z(\pi) = K((a^{1/n})^\rho)$ and $Z(\rho) = K((a^{1/n})^\pi)$ of degrees π and ρ with respect to K. We have $Z = Z(\pi)Z(\rho)$. First we prove that $\mathfrak{N}(Z^*) = \mathfrak{N}(Z(\pi)^*) \cap \mathfrak{N}(Z(\rho)^*)$. It is evident that $\mathfrak{N}(Z^*)$ is contained in the intersection of the norm groups belonging to the fields $Z(\pi)$ and $Z(\rho)$, because these fields are contained in Z. Now let a be an element contained in the intersection of the norm groups. We have

$$a = \mathfrak{N}(Z(\pi), K \mid a(\pi)) = \mathfrak{N}(Z(\rho), K \mid a(\rho))$$

with elements $a(\pi)$, $a(\rho)$ in $Z(\pi)$, $Z(\rho)$ respectively. If we consider $a(\pi)$, $a(\rho)$ as elements of Z and form the norm then we obtain $a^\rho = \mathfrak{N}(Z, K \mid a(\pi))$ and $a^\pi = \mathfrak{N}(Z, K \mid a(\rho))$. The greatest common divisor of π and ρ equals one, therefore there exist two integers s and t such that $\pi s + \rho t = 1$. Hence $a = (a^\pi)^s (a^\rho)^t = \mathfrak{N}(Z, K \mid a(\rho))^s \mathfrak{N}(Z, K \mid a(\pi))^t = \mathfrak{N}(Z, K \mid a(\rho)^s a(\pi)^t)$.

The order of $K^*/\mathfrak{N}(Z^*) = K^*/\mathfrak{N}(Z(\pi), K \mid Z(\pi)^*) \cap \mathfrak{N}(Z(\rho), K \mid Z(\rho)^*)$ is equal to $(\pi\rho)^{m-1} = n^{m-1}$ according to theorem 8. A well known theorem of group theory states that $K^*/\mathfrak{N}(Z(\pi), K \mid Z(\pi)^*) \cong K^*/\mathfrak{N}(Z^*)/\mathfrak{N}(Z(\pi)^*)/\mathfrak{N}(Z^*)$ and $K^*/\mathfrak{N}(Z(\rho)^*) \cong K^*/\mathfrak{N}(Z^*)/\mathfrak{N}(Z(\rho)^*)/\mathfrak{N}(Z^*)$. The finiteness of the orders $[K^*:\mathfrak{N}(Z(\pi)^*)]$, $[K^*:\mathfrak{N}(Z(\rho)^*)]$ and $[K^*:\mathfrak{N}(Z^*)]$ yields the result that $\mathfrak{N}(Z(\pi)^*)/\mathfrak{N}(Z^*)$ and $\mathfrak{N}(Z(\rho)^*)/\mathfrak{N}(Z^*)$ are finite groups of indices π^{m-1} and ρ^{m-1}. Their orders are ρ^{m-1} and π^{m-1} respectively. The product of $\mathfrak{N}(Z(\pi)^*)/\mathfrak{N}(Z^*)$ and $\mathfrak{N}(Z(\rho)^*)/\mathfrak{N}(Z^*)$ in $K^*/\mathfrak{N}(Z^*)$ is direct because the orders of the components

are relatively prime. Hence $K^*/\mathfrak{N}(Z^*) = \mathfrak{N}(Z(\pi)^*)/\mathfrak{N}(Z^*) \times \mathfrak{N}(Z(\rho)^*)/\mathfrak{N}(Z^*)$. The types of both components are given by the isomorphisms

$$K^*/\mathfrak{N}(Z^*)/\mathfrak{N}(Z(\pi)^*)/\mathfrak{N}(Z^*) \cong \mathfrak{N}(Z(\rho)^*)/\mathfrak{N}(Z^*) \cong K^*/\mathfrak{N}(Z(\pi)^*)$$
$$\cong \mathfrak{Z}(\pi)_1 \times \cdots \times \mathfrak{Z}(\pi)_{m-1}$$

and

$$K^*/\mathfrak{N}(Z^*)/\mathfrak{N}(Z(\rho)^*)/\mathfrak{N}(Z^*) \cong \mathfrak{N}(Z(\pi)^*)/\mathfrak{N}(Z^*)$$
$$\cong K^*/\mathfrak{N}(Z(\rho)^*) \cong \mathfrak{Z}(\rho)_1 \times \cdots \times \mathfrak{Z}(\rho)_{m-1}.$$

By composition we obtain

$$K^*/\mathfrak{N}(Z^*) \cong (\mathfrak{Z}(\pi)_1 \times \cdots \times \mathfrak{Z}(\pi)_{m-1}) \times (\mathfrak{Z}(\rho)_1 \times \cdots \times \mathfrak{Z}(\rho)_{m-1})$$
$$\cong \mathfrak{Z}(n)_1 \times \cdots \times \mathfrak{Z}(n)_{m-1},$$

because p and r are relatively prime.

REMARK. In the proof it was essential that the norm group of the composite of two fields whose degrees are prime be equal to the intersection of the norm groups belonging to the components. In the local class field theory this fact remains valid for the composite of arbitrary abelian extensions. The following example shows that such a determination of the norm group belonging to a composite does not hold for fields of power series in more than two variables.[16]

Let $K = K'\{t_1, t_2, t_3\}$ be a field of power series in three variables with an algebraically closed field of coefficients K'. The fields $L_1 = K(t_1^{1/2})$ and $L_2 = K(t_2^{1/2})$ are cyclic of degree two, their norm groups in K^* are equal to $\mathfrak{N}(L_1^*) = \{t_1, t_2^2, t_3^2\} \times \epsilon(K)$ and $\mathfrak{N}(L_2^*) = \{t_1^2, t_2, t_3^2\} \times \epsilon(K)$. The composite $L = L_1L_2$ is of degree four over K and its norm group is given by $\{t_1^2, t_2^2, t_3^4\} \times \epsilon(K)$. It is evident that the intersection of $\mathfrak{N}(L_1^*)$ and $\mathfrak{N}(L_2^*)$ properly contains $\mathfrak{N}(L^*)$.

THEOREM 10. *To different cyclic fields Z_1, Z_2 belong different norm groups $\mathfrak{N}(Z_1^*)$ and $\mathfrak{N}(Z_2^*)$ in the common ground field K.*

PROOF. First we remark that only fields Z_1 and Z_2 which have the same degree n over K could have identical norm groups. If $[Z_1:K] \neq [Z_2:K]$ then the orders of the factor groups $K^*/\mathfrak{N}(Z_1^*)$ and $K^*/\mathfrak{N}(Z_2^*)$ are different, therefore $\mathfrak{N}(Z_1^*) \neq N(Z_2^*)$. Now let $[Z_1:K] = [Z_2:K] = n$ and $Z_1 = K(a_1^{1/n})$, $Z_2 = K(a_2^{1/n})$. If $Z_1 \neq Z_2$ then there exist no relation $a_1 = a_2^z b^n$ with an integer z and an element $b \in K^*$. This means, according to the theory of radical fields, that the subgroups $\{a_1, K^{*n}\}/K^{*n}$ and $\{a_2, K^{*n}\}/K^{*n}$ are different in K^*/K^{*n}. Now let n be a power of a single prime p. According to the proof of theorem 9 the norm groups $\mathfrak{N}(Z_1^*)$ and $\mathfrak{N}(Z_2^*)$ are generated by $\{a_1, K^{*n}\}$ and $\{a_2, K^{*n}\}$ respectively. The equality $\mathfrak{N}(Z_1^*) = \mathfrak{N}(Z_2^*)$ would lead to a relation $a_1 = a_2^z b^n$,

[16] See the developments in the next paragraph.

which certainly cannot hold for distinct fields Z_1 and Z_2. The case of arbitrary degree n is again reduced to the case where the degree is a power of a prime; this is possible because the norm group of the composite is equal to the intersection of the norm groups belonging to the components.

The reduction of the question in theorem 10 implies also a proof for an existence theorem.

THEOREM 11. *To each subgroup $\mathfrak{N}$ of K^* whose factor group $K^*/\mathfrak{N}$ is the direct product of $m - 1$ cyclic groups of order n there exists a cyclic field Z of degree n over K whose norm group coincides with the given group $\mathfrak{N}$.*

PROOF. It is again obvious that it suffices to prove the assertion of the theorem for degrees n which are powers of a single prime. As a consequence of $[K^*:\mathfrak{N}] = n^{m-1}$ and $[\mathfrak{N}:K^*] = n$ there always exist elements a in $\mathfrak{N}$ which are not equal to an n^{th} power. Form now the field $Z = K(a^{1/n})$ for any one of these elements. According to previous results it is evident that $\mathfrak{N}(Z^*) = \mathfrak{N}$. All other elements $a' \in K^*$ that lie in $\mathfrak{N}$ and are not n^{th} powers determine the same field Z according to theorem 10, because $\mathfrak{N}(Z^*) = \mathfrak{N}$. The field Z is then uniquely determined by $\mathfrak{N}$.

This proof shows that the norm group $\mathfrak{N}$ of a cyclic field coincides with the very group of numbers in K^* which determines the field Z in the theory of radical fields. For abelian fields L the norm group $\mathfrak{N}(L^*)$ is of index n^{m-1} in K, and consequently $[\mathfrak{N}(L^*):K] = n$.

THEOREM 12. *For each subgroup $\mathfrak{N}$ of K^* whose index is equal to an $(m - 1)^{\text{st}}$ power of an integer n there exists a uniquely determined abelian field L whose norm group with respect to K coincides with $\mathfrak{N}$. The Galois group of L over K is isomorphic with the factor group $\mathfrak{N}/K^*$.*

PROOF. Let $a_1, \cdots, a_r$ be a basis of N with respect to K^{*n}. The field $L = K(a_1^{1/n}, \cdots, a_r^{1/n})$ is of degree n over K and its Galois group is isomorphic with $\mathfrak{N}/K^*$ according to the theory of radical fields; obviously L is uniquely determined as the field over K containing the n^{th} roots of all elements a in $\mathfrak{N}$. Conversely $\mathfrak{N}$ consists of all elements $b \in K^*$ whose n^{th} roots lie in L. We now show that $\mathfrak{N}(L^*)$ is equal to the given group $\mathfrak{N}$. The group $\mathfrak{N}(L^*)$ contains the elements $a_1, \cdots, a_r$ and therefore $\mathfrak{N}$. According to theorem 8 its index under K^* is equal to n^{m-1}, and $\mathfrak{N}$ has according to assumption the same index. Hence $\mathfrak{N}$ and $\mathfrak{N}(L^*)$ are equal.

Further application of the theory of radical fields shows that the relation between norm groups $\mathfrak{N}$ of index n and fields L of degree n over K is a one to one correspondence which conserves the natural ordering of subgroups in K^* containing $\mathfrak{N}$ and subfields L' of L; that is to say: $\mathfrak{N}' \supseteq \mathfrak{N}$, $[K^*:\mathfrak{N}'] = n'^{m-1} \leftrightarrow L' \subseteq L$, $[L':K] = n'$.

The explicit determination of the norm group belonging to an abelian field L which is not cyclic, is possible only in quite special cases. The structure of $K^*/\mathfrak{N}(L^*)$ is given by the norm class $cl(\Phi)$ of L over K.

EXAMPLE. Let $L = K(t_1^{1/e_1}, \cdots, t_m^{1/e_m})$ and $\prod_{i=1}^{m} e_i = n$. Then t_i^{n/e_i} and

$\epsilon(K)$ form a basis of $\mathfrak{N}(L^*)$, the meromorphism can be represented by the matrix

$$\left\| \begin{matrix} n/e_1 & & \\ & \ddots & \\ & & n/e_m \end{matrix} \right\|.$$

The factor group $K^*/\mathfrak{N}(L^*)$ is the direct product of m cycles $\mathfrak{Z}(n/e_i)$.

Remark. The statements on the different indices of groups can mostly be generalized to fields $\tilde{K}$ which possess a valuation $\tilde{\mathbf{V}}$ of rank m and an algebraically closed field K' of residue classes belonging to that valuation; however, the fields $\tilde{K}$ must contain a field K which is perfect with respect to a discrete valuation $\mathbf{V}$ of rank m induced by $\tilde{\mathbf{V}}$. The valuation $\tilde{\mathbf{V}}$ need not be discrete. Such fields are for example the infinite, always abelian, extensions $\tilde{K}$ of fields K which are of the type we considered till now.[17] They are not perfect with respect to the valuation $\tilde{\mathbf{V}}$ which lies over $\mathbf{V}$. Nevertheless it is again true that each unit $\tilde{\epsilon}$ of $\tilde{K}$ is an n^{th} power. According to the definition of $K(\tilde{\epsilon})$ the unit $\tilde{\epsilon}$ is an n^{th} power in $K(\tilde{\epsilon})$ because $K(\tilde{\epsilon})$ is a finite algebraic extension of K. Hence $\tilde{\epsilon}$ is a fortiori an n^{th} power in $\tilde{K}$. If we denote the group of the values of all elements $\neq 0$ in $\tilde{K}$ by $\tilde{\Gamma} = \Gamma(\tilde{K})$, then we obtain $\tilde{K}^*/\tilde{K}^{*n} \cong \tilde{\Gamma}/n \cdot \tilde{\Gamma}$. The structure and order of $\tilde{\Gamma}/n \cdot \tilde{\Gamma}$ depend essentially on the ramifications of all finite subfields of $\tilde{K}$. The latter determine the systems of rational coefficients $\{a_i\}$ that appear at the generators α_i. The type of $\{a_i\}$ is given by a G-number of Steinitz.[18]

By similar arguments one can see that the norm group of any finite extension $\tilde{L}$ of $\tilde{K}$ is given by the relations $\tilde{K}^*/\mathfrak{N}(\tilde{L}^*) \cong \Gamma(\tilde{K})/n \cdot \Gamma(\tilde{L}) = \Gamma(\tilde{K})/\Gamma(\tilde{K}) \times \tilde{\Phi}$ and $[K^*:\mathfrak{N}(L^*)] = \mathfrak{n}(cl(\tilde{\Phi})) \leqq n^{m-1}$.[19]

One can easily construct examples of infinite fields where the norm class group belonging to a special extension $\tilde{\tilde{Z}}$ of $\tilde{K}$ is isomorphic with the Galois group of $\tilde{\tilde{Z}}$ over $\tilde{K}$: for example the field $K = K(\{t_1^{1/x}\}, \cdots, \{t_{m-1}^{1/x}\}, t_m)$ which contains all the x^{th} roots of $t_1, \cdots, t_{m-1}$ where x runs over all integers. Here $\{a_1\}, \cdots, \{a_{m-1}\}$ become isomorphic with the additive group of all rational numbers. Furthermore it is possible to construct fields in which for certain

---

[17] The Galois group of $\tilde{K}$ over K is a compact group because $\tilde{K}$ is an enumerable algebraic extension of K.

[18] E. Steinitz, *Algebraische Theorie der Körper*, Crelle vol. 137 (1910) §16.

[19] The matrices representing automorphisms and meromorphisms have in general coefficients a_{ij} with denominators. These denominators are determined by the moduli of coefficients $\{a_i\}$ of the linear forms representing the valuation group. It is not difficult to give explicit formulae with the help of G-numbers.

integers n all results of the discrete case hold and for other integers they do not; there one allows x to run over a set of integers which are relatively prime to n.

The perfect closures of the above mentioned fields $\bar{K}$ can be treated in the same fashion. The validity of the calculations depends on Hensel's criterion. The latter holds also in these perfect fields.

Finally we mention that the field of residue classes K' can be exchanged for any absolutely algebraic field over a Galois field of characteristic χ. One has ordinarily to substitute m for the exponent $m - 1$ which occurred in our previous results. In our entirely multiplicative theory it must be supposed in this case that the degree of the algebraic extension of the ground field is prime to the characteristic χ. Furthermore the theorem 6 breaks down if K' is not algebraically closed. One also has to take into consideration the G-number of K'.

6. Algebras over Maximally Perfect Fields

Let Z be a cyclic extension of degree n over K. Then the group $K^*/\mathfrak{N}(Z^*)$ is the direct product of $m - 1$ cyclic groups of order n. Therefore there exists proper division algebras D of degree n over K as centrum, namely the cross products $(a, Z/K)$ with factor sets $a \in K^*$ such that a^n is the lowest power of a power of a lying in $\mathfrak{N}(Z^*)$. The group of classes of algebras split by Z is the direct product of $m - 1$ cycles of order n.[20] Thus we have proved that there exist proper division algebras over K for all degrees n.

In the following investigations we shall develop the arithmetic and algebraic theory of division algebras over K as centrum. We shall see that these algebras have a more complicated structure than the algebras in the classical local class field theory.

Let D be a fixed normal division algebra of degree m over K. The field K is perfect with respect to m valuations

$$\mathsf{V} = \mathsf{V}(m), \mathsf{V}(m - 1), \cdots, \mathsf{V}(m - (m - i)) = \mathsf{V}(i), \cdots, \mathsf{V}(1)$$

of the ranks $m, m - 1, \cdots, i, \cdots, 1$. Each valuation $\mathsf{V}(i)$ can be extended uniquely to the algebra D. For that purpose one has only to construct valuation rings in D.

Definition. An element A_{m-i} of the division algebra D over K is called an *integral element with respect to the* $\mathsf{V}(i)$*-ring* $R(m - i)$ *in* K, if and only if its irreducible equation over K, with highest coefficient one, has coefficients in $R(m - i)$.

Lemma 4. *An element $A_{m-i} \neq 0$ of D is an integer with respect to $R(m - i)$ if and only if its reduced norm with respect to K is an element of $R(m - i)$.*

The proof of the lemma which transforms the additive definition of an integral

---

[20] For the concept of cross product and the pertaining notations see M. Deuring, *Algebren*, (Berlin 1935), Ch. V.

element into a multiplicative definition, runs just like the proof for the lemma in the case of discrete archimedian valuations.[21]

THEOREM 13. *All elements A_{m-i} of a division algebra D over K which are integral with respect to the* $\mathbf{V}(i)$*-ring* $R(m - i)$ *of* K *form a ring* $\mathfrak{M}(m - i)$ *in* D. *The elements* $A'_{m-i} \in \mathfrak{M}(m - i)$ *whose reduced norms are positive with respect to* $\mathbf{V}(i)$ *form a two-sided prime ideal* $\mathfrak{P}(m - i)$ *of* $\mathfrak{M}(m - i)$. *The residue class ring* $\mathfrak{M}(m - i)/\mathfrak{P}(m - i)$ *is a division algebra.*

PROOF. First we show that all elements $A_{m-i} \in D$ for which $\mathfrak{N}_{\text{red}}A_{m-i}$ are elements in $R(m - i)$ form a ring $\mathfrak{M}(m - i)$. Let A_{m-i} and B_{m-i} be two arbitrary elements ($\neq 0$) of the system of all integers $\mathfrak{M}(m - i)$ with respect to $R(m - i)$. Then $A_{m-i}B_{m-i}$ and $B_{m-i}A_{m-i}$ are also in $\mathfrak{M}(m - i)$. We have $\mathfrak{N}_{\text{red}}(A_{m-i}B_{m-i}) = \mathfrak{N}_{\text{red}}(A_{m-i})\mathfrak{N}_{\text{red}}(B_{m-i}) = \mathfrak{N}_{\text{red}}(B_{m-i})\mathfrak{N}_{\text{red}}(A_{m-i}) = \mathfrak{N}_{\text{red}}(B_{m-i}A_{m-i})$ in $R(m - i)$, because both $\mathfrak{N}_{\text{red}}(A_{m-i})$ and $\mathfrak{N}_{\text{red}}(B_{m-i})$ lie in $R(m - i)$. Furthermore the elements of $\mathfrak{M}(m - i)$ can be ordered, that is to say either $A_{m-i}B_{m-i}^{-1}$ and $B_{m-i}^{-1}A_{m-i}$ or $A_{m-i}^{-1}B_{m-i}$ or $B_{m-i}A_{m-i}^{-1}$ are elements in $\mathfrak{M}(m - i)$. It does not matter if we consider $A_{m-i}B_{m-i}^{-1}$ or $B_{m-i}^{-1}A_{m-i}$, because $\mathfrak{N}_{\text{red}}(A_{m-i}B_{m-i}^{-1}) = \mathfrak{N}_{\text{red}}(B_{m-i}^{-1}A_{m-i}) = \mathfrak{N}_{\text{red}}(A_{m-i})\mathfrak{N}_{\text{red}}(B_{m-i}^{-1}) = \mathfrak{N}_{\text{red}}(A_{m-i})\mathfrak{N}_{\text{red}}(B_{m-i})^{-1}$. If $A_{m-i}B_{m-i}^{-1}$ is already an element of $\mathfrak{M}(m - i)$ then there is nothing to be proved. Assume now that $A_{m-i}B_{m-i}^{-1}$ is not contained in $\mathfrak{M}(m - i)$. Then its reduced norm is not an element of $R(m - i)$. The ring $R(m - i)$ is a commutative valuation ring; therefore the inverse of $\mathfrak{N}_{\text{red}}(A_{m-i}B_{m-i}^{-1})$ lies in $R(m - i)$. This inverse is the reduced norm of $A_{m-i}^{-1}B_{m-i}$ or $B_{m-i}A_{m-i}^{-1}$. According to the definition of $\mathfrak{M}(m - i)$, the elements $B_{m-i}A_{m-i}^{-1}$ and $A_{m-i}^{-1}B_{m-i}$ lie in $\mathfrak{M}(m - i)$ as asserted.

Now we show that the sum of A_{m-i} and B_{m-i} lies also in $\mathfrak{M}(m - i)$. In virtue of the ordering in $\mathfrak{M}(m - i)$ we can assume without loss of generality that $A_{m-i}B_{m-i}^{-1} \in \mathfrak{M}(m - i)$. Therefore the minimal equation $f(x) = 0$ of $A_{m-i}B_{m-i}^{-1}$ has integral coefficients with respect to $R(m - i)$. The polynomial $f(x - 1)$ which furnishes the minimal equation of $A_{m-i}B_{m-i}^{-1} + 1$ has also integral coefficients; according to the additive definition of the integers in $\mathfrak{M}(m - i)$ the element $A_{m-i}B_{m-i}^{-1} + 1$ lies in $\mathfrak{M}(m - i)$. Hence

$$(A_{m-i}B_{m-i}^{-1} + 1)B_{m-i} = A_{m-i} + B_{m-i}$$

in $\mathfrak{M}(m - i)$. The ring $\mathfrak{M}(m - i)$ contains the ring $R(m - i)$ and is obviously maximal by virtue of its definition.

All elements $A'_{m-i} \in \mathfrak{M}(m - i)$ whose reduced norm is positive with respect to $\mathbf{V}(i)$, i.e. $\mathbf{v}(i)(\mathfrak{N}_{\text{red}}A'_{m-i}) > 0$, form a two-sided ideal $\mathfrak{P}(m - i)$ in $\mathfrak{M}(m - i)$. This is easily seen by considering the norms of $A'_{m-i} + B'_{m-i}$, $A_{m-i}A'_{m-i}$ and $A'_{m-i}A_{m-i}\cdot(A'_{m-i}, B'_{m-i} \in \mathfrak{P}(m - i), A_{m-i} \in \mathfrak{M}(m - i))$.

In order to prove that $\mathfrak{P}(m - i)$ is a prime ideal we must show that

[21] The model of these proofs will be found in the paper of H. Hasse, *Über p-adische Schiefkörper und ihre Bedeutung für die Arithmetik hyperkomplexer Zahlsysteme*, Math. Annalen vol. 104 (1931), §2.

$\mathfrak{M}(m-i)/\mathfrak{P}(m-i)$ is a simple algebra. Here we can prove more; the residue ring is a division algebra. All residue classes of $\mathfrak{M}(m-i)/\mathfrak{P}(m-i)$ can be represented by units η of $\mathfrak{M}(m-i)$. Let η be an arbitrary representative of some residue class $\neq 0$; the unit η^{-1} fulfills the congruence $\eta\eta^{-1} = \eta^{-1}\eta \equiv 1$ (mod $\mathfrak{P}(m-i)$), and therefore its uniquely determined residue class η^{-1} mod $\mathfrak{P}(m-i)$ is the inverse of η mod $\mathfrak{P}(m-i)$.

The algebra $\mathfrak{M}(m-i)/\mathfrak{P}(m-i)$ is in general not commutative, and moreover the field $R(m-i)/\mathfrak{p}(m-i)$ need not be the centrum of $\mathfrak{M}(m-i)/\mathfrak{P}(m-i)$ as examples we shall mention later will prove.

In the case of the valuation $\mathbf{V}(1)$ of rank one all ideals of $\mathfrak{M}(m-1)$ are powers of the principal prime ideal $\mathfrak{P}(m-1)$. In particular the prime ideal $\mathfrak{p}(m-1)$ of $R(m-1)$ becomes an E^{th} power of $\mathfrak{P}(m-1)$. Let P be any element of $\mathfrak{P}(m-1)$ which is exactly divisible by $\mathfrak{P}(m-1)$. The element P generates in D a field $K(P)$ which is at most of degree n over K. Its prime ideal $\mathfrak{P}$ evidently generates in $\mathfrak{M}(m-1)$ the prime ideal $\mathfrak{P}(m-1)$. Therefore we have $\mathfrak{p}(m-1) = \mathfrak{P}^E$, and as a consequence of $[K(P):K] \leqq n$ we obtain $E \leqq n$. We shall show by examples that E actually can be equal to 1 for proper division algebras of degree n. Let the rank of $\mathfrak{M}(m-1)/\mathfrak{P}(m-1)$ over $R(m-1)/\mathfrak{p}(m-1)$ be F. According to the theory of discrete valuations we have $EF = n^2$.[22] Therefore F is greater or equal to n.

Let $\bar{C}$ be the centrum of $\mathfrak{M}(m-1)\mathfrak{P}(m-1)$; then $[\mathfrak{M}(m-1)/\mathfrak{P}(m-1):\bar{C}] = r^2$, $[\bar{C}:R(m-1)/\mathfrak{p}(m-1)] = s$ and $r^2 s = F$. Any maximally commutative subfield $\bar{H}$ of $\mathfrak{M}(m-1)/\mathfrak{P}(m-1)$ is of degree rs over $R(m-1)/\mathfrak{p}(m-1)$. Let $\bar{\Theta}$ be a primitive element of $\bar{H}$ and let $\bar{f}(x) = x^{rs} + \bar{a}_1 x^{rs-1} + \cdots + \bar{a}_{rs} = 0$ be its irreducible equation in $R(m-1)/\mathfrak{p}(m-1)$. The residue class $\bar{\Theta}$ may be represented by some unit $\Theta \in \mathfrak{M}(m-1)$. The unit Θ generates a subfield $K(\Theta)$ of D. The irreducible polynomial $f(x)$ belonging to Θ is congruent to $\bar{f}(x)$ modulo $\mathfrak{p}(m-1)$. Hence the degree $[K(\Theta):K]$ is equal to $rs = [\bar{H}:R(m-1)/\mathfrak{p}(m-1)]$ according to Hensel's criterion. Therefore we have $rs \leqq n$. The field $K(\Theta)$ is unramified with respect to $\mathfrak{p}(m-1)$ because Θ and its conjugates generate, if properly selected, an order in $K(\Theta)$ which has $\bar{H}$ as its system of residue classes modulo $\mathfrak{p}(m-1)$. Other representatives Θ' obviously generate other fields $K(\Theta')$ which are isomorphic with $K(\Theta)$ because they are all isomorphic with fields of power series in one variable with fields $\bar{H}$ of coefficients that are all isomorphic with $\bar{H}$. Splitting fields $\bar{H}_1$ and $\bar{H}_2$ of $\mathfrak{M}(m-1)/\mathfrak{P}(m-1)$ which are not isomorphic determine subfields H_1 and H_2 of D which are not isomorphic.

In every case $rs = n$ because there exist units in $\mathfrak{M}(m-1)$ whose irreducible equations are exactly of degree n. The residue class of such a unit mod $\mathfrak{P}(m-1)$ has also the degree n as a consequence of Hensel's criterion. We have proved the following theorem.

[22] Cf. Deuring's Report ch. VI, §12.

THEOREM 14. *Each division algebra D of degree n over K contains maximally commutative subfields L which are unramified with respect to the valuation* $\mathbf{V}(1)$ *of the centrum K.*

Assume now that $E = n$ for a division algebra D. Then any element P which is exactly divisible by $\mathfrak{P}(m-1)$ defines a maximally commutative subfield $K(P)$ of D. According to the construction the prime ideal $\mathfrak{p}(m-1)$ is also totally ramified in $K(P)$. Because of theorem 4 the field $K(P)$ is cyclic over K and therefore equal to a radical field $K(t_1^{a_1}, \cdots t_m^{a_m})^{1/n})$; lemma 3 yields that a_m is prime to n. Hence there exists an integer x prime to n such that $a_m x \equiv 1 \pmod n$. The radical $(t_1^{a_1 x} \cdots t_{m-1}^{a_{m-1}x} t_m)^{1/n}$ also generates the field $K(P)$. According to the theory of cyclic algebras there exists a factor set $b' = t_1^{c_1} \cdots t_m^{c_m}$ in K^* such that $D \sim (b', K(P)/K)$. Now the element $t_1^{a_1 x} \cdots t_m$ is the norm of $(t_1^{a_1 x} \cdots t_m)^{1/n}$; therefore we can substitute for b' the element $b'(t_1^{a_1 x} \cdots t_m)^{-c_m} = b$ and we obtain $D \sim (b, K(P)/K)$. We have proved the following theorem.

THEOREM 15. *Each division algebra D of degree n which is totally ramified with respect to the valuation* $\mathbf{V}(1)$ *of the centrum K can be represented as a cross product* $(t_1^{b_1} \cdots t_{m-1}^{b_{m-1}}, K((t_1^{a_1} \cdots t_{m-1}^{a_{m-1}} t_m)^{1/n})/K)$.

REMARK. These totally ramified algebras have an unramified cyclic splitting field $K((t_1^{b_1} \cdots t_{m-1}^{b_{m-1}})^{1/n})$; as one easily observes by direct computation $t_1^{a_1} \cdots t_{m-1}^{a_{m-1}} t_m$ is the factor set with respect to that unramified representation.

The field K is also perfect with respect to the valuation $\mathbf{V}(m-1)$. The system of residue classes $\mathfrak{M}(1)/\mathfrak{P}(1)$ is a commutative field over $R(1)/\mathfrak{p}(1) = K(1) \cong K'\{t_1\}$. In order to prove this assertion we have to show that there exist no division algebras $D(1)$ over $K(1)$. According to theorem 6 there exists only one extension $Z(1) = K(1)(t_1^{1/n})$ of degree n over $K(1)$. Any maximally commutative subfield of $D(1)$ is therefore isomorphic with $Z(1)$; hence $D(1) \sim (b(1), Z(1)/K(1))$ with suitable $b(1) \in K(1)^*$. But according to theorem 8 the element $b(1)$ is norm of an element in $Z(1)$, and therefore $D(1) \sim 1$.

We consider now the special class of algebras D of degree n over K for which $[\mathfrak{M}(1)/\mathfrak{P}(1):K(1)] = n$, that is to say for which $\mathfrak{M}(1)/\mathfrak{P}(1) \cong Z(1)$. According to Hensel's criterion the algebra D contains a cyclic field $K(\Theta)$ of degree n over K whose residue system modulo $\mathfrak{p}(1)$ is a field isomorphic with $Z(1)$. All fields in D which are unramified with respect to $\mathfrak{p}(1)$ are isomorphic with $K(\Theta)$ because $Z(1)$ is uniquely determined. They are isomorphic with $K(t_1^{1/n})$ as a consequence of Hilbert's ramification theory, the field $K(t_1^{1/n})$ being its own field of inertia with respect to $\mathfrak{p}(1)$. These considerations show that $K(t_1^{1/n})$ is a splitting field for all division algebras D of degree n whose residue class degrees with respect to $\mathfrak{p}(1)$ are equal to n. An obvious consequence is the following theorem.

THEOREM 16. *All algebras of degree n whose residue class degrees with respect to the valuation* $\mathbf{V}(m-1)$ *of K are equal to n form a group which is the direct product of $m-1$ cycles of order n, the field $K(t_1^{1/n})$ is the common unramified cyclic splitting field.*

REMARK 1. This theorem enables us to determine the structure of the system of all classes of algebras over a field of power series in two variables. We first show that all division algebras of degree n form a cyclic group of order n. For $m = 2$ the prime ideal $\mathfrak{p}(1)$ is a principal ideal. Therefore we obtain the numbers E and F as introduced previously for division algebras. The number F is equal to $[\mathfrak{M}(1)/\mathfrak{P}(1):K(1)]$. This degree is at most equal to n, because the degree of the irreducible equation of any unit representing a residue class of the cyclic field $\mathfrak{M}(1)/\mathfrak{P}(1)$ is at most equal to n. Furthermore $E \leqq n$ and $EF = n^2$. Hence $E = F = n$. Theorem 16 asserts that each division algebra of degree n is split by $K(t_1^{1/n})$. Consequently all classes of algebras representable by division algebras of degree n'/n are exhausted by the group of all algebras which are split by $K(t_1^{1/n})$. The latter group is cyclic of order n as theorem 9 asserts.

A word for word application of the procedure in local class field theory shows that the group of all normal algebras over K is isomorphic with the additive group of all rational fractions modulo 1. Consequently all statements of the classical local class field theory hold for fields of power series in two variables with an algebraically closed field of coefficients.[23]

REMARK 2. For a division algebra $D = (t_1^{a_1} \cdots t_{m-1}^{a_{m-1}}, K((t_1^{1/n})/K))$ we certainly have $[\mathfrak{M}(m-1)/\mathfrak{P}(m-1):R(m-1)/\mathfrak{p}(m-1)] = n^2$ if m is greater than two. This is proved by considering the hypercomplex order generated by the units $u^i t_1^j$ $(i, j = 1, 2, \cdots, n)$, with respect to $\mathfrak{p}(m-1)$, over $R(m-1)$. One easily sees that these units are also linearly independent modulo $\mathfrak{P}(m-1)$. Therefore $E = 1$ and $F = n^2$. Furthermore one observes that

$$\mathfrak{M}(m-1)/\mathfrak{P}(m-1)$$

is a division algebra with the centrum $K(m-1)$. This fact can also be stated in the following way: D is isomorphic with the system of power series in one variable t_m with coefficients in the division algebra $\mathfrak{M}(m-1)/\mathfrak{P}(m-1)$, where the powers of t_m are permutable with the coefficients.

REMARK 3. If $n = p$ then there exist only two possibilities for

$$\mathfrak{M}(m-1)/\mathfrak{P}(m-1).$$

The algebra of residue classes is either a division algebra of degree p over $K(m-1)$ and $E = 1$, or it is a field of degree p over $K(m-1)$ and $E = F = p$. At the end of this paragraph we shall illustrate by examples that for general n there arise much more complicated situations.

THEOREM 17. *The number of classes of algebras representable by division algebras of degree n'/n is greater than n^{m-1} if the reference field K is a field of power series of more than two variables.*

PROOF. The number of classes split by any cyclic field of degree n over K

[23] For example Cl. Chevalley, *La théorie du symbole de restes normiques*, Crelle vol. 169 (1932).

is equal to n^{m-1}. In order to prove the assertion of the theorem it is only necessary to construct a division algebra with a cyclic splitting field which is not split by another cyclic field that will also be constructed. Then the groups of algebras generated by the cross products belonging to the two cyclic fields do not coincide entirely, and hence the number of classes of degree n is greater than n^{m-1}.

Let $Z_1 = K(t_2^{1/n})$, $Z_2 = K(t_3^{1/n})$ and $D = (t_1, Z_1/K)$ be two cyclic fields and a division algebra over K. We show that $D \times Z_2 \not\sim Z_2$. This assertion is equiavlent to the statement that t_1 is not the norm of an element in Z_1Z_2 if it is considered as an element of Z_2. We need only observe that t_1 is not contained in the norm group $\mathfrak{N}((Z_1Z_2)^*) = \{t_1^n, t_2, t_3^{1/n}, \cdots, t_m^n\} \times \epsilon(Z_2)$.

THEOREM 18. *The group of algebras split by an arbitrary field L over K is finite.*

PROOF. Let $A = (a_{\sigma,\tau}, L/K)$ be an arbitrary algebra split by L. ($\sigma, \tau, \cdots$ the elements of the Galois group of L over K, $[L:L] = n$). The exponent of A is at most equal to n; hence $a_{\sigma,\tau}^n = c_\sigma c_\tau^\sigma / c_{\sigma\tau}$ with suitable elements $c_\sigma \in L^*$. According to a theorem of R. Brauer, we have $A \sim (a'_{\sigma',\tau'}, L'(a_{\sigma,\tau})/K)$ where $L'(a_{\sigma,\tau})$ is a normal field over K which contains the field $L(c_\sigma^{1/n}, c_\tau^{1/n}, \cdots)$ and $\sigma', \tau', \cdots$ elements in the Galois group of $L'(a_{\sigma,\tau})$ over K.[24] Under the assumptions on K the field $L'(a_{\sigma,\tau})$ is equal to $L(c_\sigma^{1/n}, \cdots)$. The factor set $a'_{\sigma',\tau'}$ consists of n^{th} roots of unity. Obviously the group of all possible factor sets $\{a'_{\sigma',\tau'}\}$ consisting of roots of unity is finite. Hence the subgroup of all algebras which belong already to L and to the same field $L'(a_{\sigma,\tau})$ is finite.

For a fixed factor set $a_{\sigma,\tau}$ the degree $[L'(a_{\sigma,\tau}):K]$ is a divisor of the number $n \cdot n^n = n^{n+1}$. Thus the degrees of the different fields $L'(a_{\sigma,\tau})$ belonging to different factor sets of L are uniformly bounded by n. Now the number of fields L' over K whose degrees are divisors of n^{n+1} is bounded according to theorems 6 and 7. Both statements on the finiteness of possibilities affirm the assertion of the theorem.

THEOREM 18′. *The number of classes of normal algebras of degree n'/n over K is finite.*

PROOF. Each class can be represented by a division algebra. The latter are always split by abelian fields. Since there exists only a finite number of fields of degree n'/n, application of theorem 18 to all these fields proves the finiteness.

We return again to the algebras A of the form $(a, K(t_1^{1/n})/K)$ and ask under what conditions an arbitrary field L of degree n over K is a splitting field of such an algebra.

For this purpose we consider K as a maximally perfect field with respect to the valuation $\mathsf{V}(m-1)$ and the associated field of residue classes $K(1) \cong K'\{t_1\}$. The field of inertia $T(m)_1$ of L with respect to $\mathsf{V}(m-1)$ is equal to $K(t_1^{1/h})$,

[24] R. Brauer, *Über die Konstruktion der Schiefkörper, die von endlichem Rang in bezug auf ein gegebenes Zentrum sind*, Crelle vol. 168 (1932).

as renewed application of Hensel's criterion shows. The degree of L over $T(m)_1$ is equal to $\mathfrak{n}(\bar{\mathrm{E}}(\Gamma/\Delta_1))$, where $\bar{\mathrm{E}}(\Gamma/\Delta_1)$ denotes a representing meromorphism belonging to the ramification class of L over K with respect to the valuation $\mathbf{V}(m-1)$.

The group $K^*/\mathfrak{N}((K(t_1^{1/n}))^*)$ is the direct product of $m-1$ cyclic groups. The elements $t_2, \cdots t_m$ are surely not norms of elements in $K(t_1^{1/n})$. They obviously represent a system of generators of the group $\mathfrak{N}((K(t_1^{1/n}))^*)$. Furthermore any unit $\epsilon(K)_{m-1}$ of K with respect to the valuation $\mathbf{V}(m-1)$ is the norm of a unit in $K(t_1^{1/n})$. We have

$$K^*/\mathfrak{N}((K(t_1^{1/n}))^*) = \{t_2, \cdots, t_m\}/\{t_2^n, \cdots, t_m^n\} \times \epsilon(K)_{m-1}/\mathfrak{N}(\epsilon(K(t_1^{1/n}))_{m-1});$$

therefore $[\epsilon(K)_{m-1}:\mathfrak{N}(\epsilon(K(t_1^{1/n}))_{m-1})] = 1$. This result for the units $\epsilon(K)_{m-1}$ of K with respect to the valuation $\mathbf{V}(m-1)$ holds for any field $\tilde{K}$ which is a field of power series in m variables, and any finite algebraic extension of $\tilde{K}$ which is generated by the n^{th} root of an element in $\tilde{K}$ whose value with respect to $\tilde{\mathbf{V}}(m)$ is a generator of the isolated subgroup $\tilde{\Delta}_1$ of $\tilde{\Gamma}$.

For the solution of our problem we have to consider the algebra $A \times L \sim (a, L\cdot K(t_1^{1/n})/L, \sigma(L))$, where $\sigma(L)$ denotes the least power of the generator σ of the Galois group of $K(t_1^{1/n})$ over K which generates the Galois group of $L\cdot K(t_1^{1/n})$ over L. Furthermore the degree $[L\cdot K(t_1^{1/n}):L]$ is equal to n/h, where $h = [L \cap K(t_1^{1/n}):K]$. The field $L \cap K(t_1^{1/n})$ is the field of inertia $T(m)_1$ which belongs to $\mathbf{V}(m-1)$ in L, and we therefore have $h = n\cdot\mathfrak{n}(\bar{\mathrm{E}}(\Gamma/\Delta_1))^{-1}$.

Let the valuation generated by $\mathbf{V}(m-1)$ in L be denoted by $\mathbf{V}(m-1)_L$. According to an observation on the norm groups belonging to unramified extensions with respect to the valuation $\mathbf{V}(m-1)$ the element a considered as an element of L is the norm of an element in $L\cdot K(t_1^{1/n})$ if and only if its value $\mathbf{v}(m-1)_L(a)$ belongs to the group of values generated by the norms of element in $L\cdot K(t_1^{1/n})$. Because that field is unramified over L with respect to $\mathbf{V}(m-1)$ the necessary and sufficient condition that a be a norm is given by the congruence

$$\mathbf{v}(m-1)_L(a) \equiv 0 \ (\mathrm{mod}\ n/h = \mathfrak{n}(\bar{\mathrm{E}}(\Gamma/\Delta_1))).$$

Hence we obtain the following theorem.

THEOREM 19. *An algebra* $A = (a, K(t_1^{1/n})/K)$ *is split by a field* L *of degree* n *over* K *if and only if*

$$\mathbf{v}(m-1)_L(a) \equiv 0 \ (\mathrm{mod}\ \mathfrak{n}(\bar{\mathrm{E}}(\Gamma/\Delta_1))).$$

REMARK. Theorem 19 asserts anew that in fields of power series of two variables the local class field theory holds. The valuation $\mathbf{V}(1)$ is of rank one, and $\mathbf{v}(1)_L(a) = e\cdot\mathbf{v}(1)(a)$. Furthermore $f = h$. Therefore the condition for a splitting field becomes in this case $\mathbf{v}(1)_L(a) = e\cdot\mathbf{v}(1)(a) \equiv 0$ (mod $n/f = e$). It is always fulfilled for any field L of degree n over K. Therefore the group of algebras split by $K(t_1^{1/n})$ coincides with the group of algebras belonging to any field L of degree n over K. This is the law of reciprocity from which all other theorems of local class field theory are more or less consequences.

Finally we wish to discuss two examples which illustrate the difference between fields of power series in more than two variables and fields of power series in two variables, for which as we have seen, the usual local class field theory holds.

EXAMPLE 1. Let $K = K'\{t_1, t_2, t_3, t_4\}$ be a field of power series in four variables. Now consider the division algebras $D_1 = (t_1, K(t_2^{1/p})/K)$ and $D_2 = (t_4, K(t_3^{1/p})/K)$, with the defining relations $u_\sigma^{-1}(t_2^{1/p})u_\sigma = \zeta_p(t_2^{1/p})$, $u_\sigma^p = t_1$ and $u_\tau^{-1}(t_3^{1/p})u_\tau = \zeta_p'(t_3^{1/p})$, $u_\tau^p = t_4$ respectively. (ζ_p and ζ_p' denote suitable p^{th} roots of unity.) Then consider the algebra A determined by the relations $u_\sigma u_\tau = u_\tau u_\sigma$, $u_\sigma^p = t_1$, $u_\tau^p = t_4$; u_τ leaves $t_2^{1/p}$ invariant and u_σ leaves $t_4^{1/p}$ invariant. The coefficients of the sums $\sum_{i,j=1}^{p} a_{ij}\, u_\sigma^i\, u_\tau^j$ shall be taken in the field $K(t_2^{1/p}, t_3^{1/p})$ which is unramified with respect to $\mathbf{V}(1)$. The algebra A is normal over K and has the degree p^2. It contains two algebras which are isomorphic with D_1 and D_2. Both have the centrum K, therefore $A \sim D_1 \times D_2$. The elements $t_3^{1/p}$, $t_2^{1/p}$, u_σ, u_τ and their products generate over $R(3)$ an order of rank p^4. This order $\mathfrak{O}$ is contained in at least one maximal order $\mathfrak{M}$ of A.[25] Denote by $\mathfrak{P}$ the uniquely determined two-sided ideal of $\mathfrak{M}$ with respect to the valuation $\mathbf{V}(1)$. We then get the following relations: $\mathfrak{p} = \mathfrak{P}^E$, $[\mathfrak{M}/\mathfrak{P}:K(3)] = F$, $EF = p^4$ and $E = p^\lambda$, $F = p^\mu$ with $\lambda + \mu = 4$. We have $\lambda \geqq 1$ and $\mu \leqq 3$, since A contains the field $K(u_\sigma) \cong K(t_4^{1/p})$ which is totally ramified. The system of residue classes $\mathfrak{M}/\mathfrak{P}$ is a simple algebra over $K(3)$. It contains the ring $\mathfrak{O}/\mathfrak{P}$. Let the residue classes of $t_3^{1/p}$, $t_2^{1/p}$ and u_σ be denoted by $\bar{t}_3^{1/p}$, $\bar{t}_2^{1/p}$ and $\bar{u}_\sigma$. They are not equal to 0 in $K(3)$; furthermore there occur no new relations between them. Hence $\mathfrak{M}/\mathfrak{P}$ contains the algebra $\bar{B} = K(3)(\bar{t}_2^{1/p}, \bar{t}_3^{1/p}, \bar{u}_\sigma)$. The system $\bar{B}$ is a cyclic algebra of rank p^2 over $K(3)(\bar{t}_3^{1/p})$, namely $\bar{B} = (\bar{t}_1, K(3)(\bar{t}_2^{1/p}, \bar{t}_3^{1/p})/K(3)(\bar{t}_3^{1/p}))$; it is a division algebra. Its rank over $K(3)$ is equal to p^3. Therefore the system $\mathfrak{M}/\mathfrak{P}$ which is at most of rank p^3 over $K(3)$ must coincide with $\bar{B}$. We obtain $\mu = 3$ and $\lambda = 1$. The algebra A does not split off a matrix algebra of degree p because $\mathfrak{M}/\mathfrak{P}$ is a division algebra.

By this example we have shown:

(i) $A \sim D_1 \times D_2$ is a division algebra with centrum K; it has the degree p^2 and is of exponent p.

(ii) $\mathfrak{M}/\mathfrak{P}$ is a division algebra of rank p^3 whose centrum properly contains $K(3)$.

(iii) A has more than one type of abelian splitting fields which are unramified with respect to $\mathbf{V}(3)$. For example $K(t_2^{1/p}, t_3^{1/p})$ and $K(t_1^{1/p}, t_3^{1/p})$.

EXAMPLE 2. Let $K = K'\{t_1, t_2, t_3\}$ be the reference field. The algebra $D = (t_2 t_3^p, K(t_1^{1/\pi})/K)$ is a division algebra of degree $p^2 = \pi$ and exponent p^2 over K because the $p^{2\text{th}}$ power of $t_2 t_3^p$ is the least which lies in the norm group of $K(t_1^{1/\pi})$ in K, the latter is equal to $\{t_1, t_2^\pi, t_3^\pi\} \times \epsilon(K)$.

---

[25] The order $\mathfrak{O}$ is contained in at least one maximal order $\mathfrak{M}$ because for valuations of rank one maximal orders can be characterized by the postulate that they have a minimal discriminant. It is necessary to have at this step the existence of such a maximal order for we do not yet know that A is a division algebra.

We now show that no field $K((t_1^{a_1}t_2^{a_2}t_3)^{1/\pi})$ is a splitting field of D. Denote the field $K((t_1^{a_1}t_2^{a_2}t_3)^{1/\pi}, t_1^{1/\pi})$ by L. Then $D \times K((t_1^{a_1}t_2^{a_2}t_3)^{1/\pi}) \sim (t_2t_3^p, L/K((t_1^{a_1}t_2^{a_2}t_3)^{1/\pi}))$, because $K(t_1^{1/\pi}) \cap K((t_1^{a_1}t_2^{a_2}t_3)^{1/\pi}) = K$. The group of all elements unequal to 0 in $K((t_1^{a_1}t_2^{a_2}t_3)^{1/\pi})$ is given by $\{(t_1^{a_1}t_2^{a_2}t_3)^{1/\pi}, t_1, t_2\} \times \epsilon(K((t_1^{a_1}t_2^{a_2}t_3)^{1/\pi}))$, because the norms of the generators $(t_1^{a_1}t_2^{a_2}t_3)^{1/\pi}$, t_1 and t_2 generate exactly the norm group of $K((t_1^{a_1}t_2^{a_2}t_3)^{1/\pi})$ with respect to K. The group L^* is equal to

$$\{(t_1^{a_1}t_2^{a_2}t_3)^{1/\pi}, t_1^{1/\pi}, t_2\} \times \epsilon(L),$$

as a computation similar to the one we just made shows. Now $t_2t_3^p$ is not contained in the norm group of L in $K((t_1^{a_1}t_2^{a_2}t_3)^{1/\pi})$. If $t_2t_3^p$ were an element of the norm group we could establish a relation $t_2t_3^p = (t_1^{a_1}t_2^{a_2}t_3)^x t_1^y t_2^{\pi z}$ with integers x, y and z. If we assort the powers of t_1, t_2, t_3 and compare both sides, we obtain the linear equations $a_1x + y = 0$, $a_2x + p^2z = 1$ and $x = p$. Substitution of $x = p$ in $a_2x + p^2z = 1$ leads to the relation $(a_2 + pz)p = 1$. This equation has no integral solution z; therefore $t_2t_3^p$ is not an element of the norm group. The same is true for the p^{th} power of $t_2t_3^p$. Hence the algebra $D \times K((t_1^{a_1}t_2^{a_2}t_3)^{1/\pi})$ is a division algebra over $K((t_1^{a_1}t_2^{a_2}t_3)^{1/\pi})$, and therefore a field of the type $K((t_1^{a_1}t_2^{a_2}t_3)^{1/\pi})$ never can be a splitting field of D.

We now prove that the prime ideal $\mathfrak{p}(2)$ of K never can be a $p^{2\text{th}}$ power of a prime ideal $\mathfrak{P}(2)$ of D. Assume that $\mathfrak{p}(2) = \mathfrak{P}(2)^\pi$; then any element P of D which is exactly divisible by $\mathfrak{P}(2)$ generates a field $K(P)$ in D which is of degree p^2 over K and totally ramified with respect to $\mathfrak{p}(2)$. According to theorem 4 the field $K(P)$ must be cyclic and therefore be equal to some field $K((t_1^{a_1}t_2^{a_2}t_3)^{1/\pi})$. The latter cannot be a maximally commutative subfield of D. Hence $\mathfrak{p}(2)$ cannot become a $p^{2\text{th}}$ power of the two-sided prime ideal $\mathfrak{P}(2)$ of D. But we have $p(2) = \mathfrak{P}(2)^p$ because D contains the field $K(u)$ generated by the defining operator u whose $p^{2\text{th}}$ power is equal to $t_2t_3^p$. For that reason the degree of ramification is equal to p, and therefore the degree of ramification belonging to $\mathfrak{p}(2)$ in D must also be equal to p. The algebra D is an example of a cyclic algebra of degree and exponent p^2 with the ramification degree p and the residue class degree p^3 with respect to $\mathbf{V}(3)$.

THE INSTITUTE FOR ADVANCED STUDY.

ANNALS OF MATHEMATICS
Vol. 38, No. 3, July, 1937

ON THE GROUP-DEFINING RELATIONS (2, 3, 7; p)

BY ABRAHAM SINKOV

(Received December 10, 1936)

1. **Introduction.** The defining relations $S^3 = T^2 = (ST)^7 = (S^{-1}T^{-1}ST)^p = 1$, which are designated (2, 3, 7; p), have been studied for all values of p up to and including seven.[1] In each case, the four conditions have been shown to be either incompatible or else sufficient to define a single finite group. That such a statement is not generally true follows from the fact that it is possible to exhibit a pair of generators of the simple group of order 1092, and a pair of generators of the simple group of order 9828, such that both pairs[2] satisfy the relations (2, 3, 7; 13). Further it would appear from the work of Brahana on the simple group of order 504 that there is more than one group[3] satisfying the relations (2, 3, 7; 9). Indeed it is quite probable that this same statement is true for every value of $p > 7$. The purpose of the present paper is to obtain some new information about the general relations (2, 3, 7; p) and incidentally to determine the abstract definitions of certain groups which have not previously been defined in this manner.

2. **General Relations between P and Q.** Under the transformation

$$P = (ST)^{-1}, \qquad S = P^2Q,$$
$$Q = (ST)^2S, \qquad T = P^3Q,$$

the defining relations (2, 3, 7; p) are equivalent[4] to $P^7 = Q^p = (QP^2)^3 = (QP^3)^2 = 1$. The latter definition is more convenient to work with, and will be used in this paper in preference to the definition in terms of S and T. Certain consequences of these relations, independent of the period of P, have been established by the author in the paper referred to in footnote 2. We restate them here, without proof:

$$QP^2Q = PQP, \tag{1}$$

$$(P^2QPQ^\alpha)^2 = (PQP^2Q^{\alpha+1})^2 = 1, \tag{2}$$

---

[1] H. R. Brahana, American Journal of Mathematics, vol. 50 (1928) pp. 350–4. This paper will hereafter be referred to as "Certain Perfect Groups."
A. Sinkov, Bulletin of the American Mathematical Society, vol. 41 (1935) p. 239.

[2] A. Sinkov, "Necessary and Sufficient Conditions for Generating Certain Simple Groups," Am. Journ. of Math., vol. 49 (1937) p. 71.
H. R. Brahana, Annals of Mathematics, vol. 31 (1930) p. 532.

[3] "Certain Perfect Groups," p. 354.

[4] "Certain Perfect Groups," p. 349.

$$(3) \qquad QP^{-1} = PQP^4Q,$$

$$(4) \qquad (QP^4)^3 = 1,$$

$$(5) \qquad (Q^2P^2)^3 = 1.$$

From these we derive two additional relations which are true only when $P^7 = 1$:

$$\begin{aligned} Q^{-2}P^4 &= Q^{-1}\cdot P^3Q = Q^{-1}P^4\cdot P^6Q \\ &= P^3QP^6Q = P^4\cdot P^6QP^2\cdot P^4Q \\ &= P^4QP\cdot Q^{-1}P^4\cdot Q = P^4QP^4Q^2, \\ Q^{-3}P^4 &= Q^{-1}P^4\cdot QP^4Q^2 = P^3Q^2P^4Q^2 \\ &= P^3Q^3\cdot Q^{-1}P^4Q^{-1}\cdot Q^3 = P^3Q^3P^3Q^3, \end{aligned}$$

whence (6) $(Q^3P^3)^3 = 1$;

$$\begin{aligned} (Q^2P^5Q^2P^3)^2 &= Q^2P^5\cdot QP^4Q\cdot P^5Q^2P^3 \\ &= Q^2P^4QP^4Q^2P^3 = QP^3QP^3, \end{aligned}$$

$$(7) \qquad (Q^2P^5Q^2P^3)^2 = 1.$$

3. **On the Operator** Q^2P^5. It has already been stated that, for $p > 7$, the four conditions (2, 3, 7; p) seem insufficient to determine a unique group. If that be true, then there are certain combinations of the generators whose periods cannot be fixed by those four conditions alone. For example, if we consider the operators $Q^\alpha P^\beta$ arranged in the ascending order of the numbers $10\,\alpha + \beta$, then it is possible to show that the period of every operator up to and including Q^2P^4 is fixed by the initial relations. This is not true however for Q^2P^5, and an attempt will now be made to indicate some limitations that can be placed upon its possible periods.

From the relation (7) just established,

$$\begin{aligned} Q^2P^5Q^2P^3 &= P^4Q^{-2}P^2Q^{-2}, \\ (Q^2P^5)^2 &= P^4Q^{-2}(Q^2P^5)^{-2}Q^2, \end{aligned}$$

$$(8) \qquad (Q^2P^5)^2Q^{-2}(Q^2P^5)^2 = P^4Q^{-2},$$

$$(9) \qquad (Q^2P^5)^2Q^{-2}(Q^2P^5)^3 = P^2,$$

$$(10) \qquad (Q^2P^5)^3Q^{-2}(Q^2P^5)^3 = Q^2.$$

If the period of Q^2P^5 were two, then (8) would lead to $P^4 = 1$, which is impossible.

Suppose $(Q^2P^5)^3 = 1$. Then it results from (10) that $Q^4 = 1$. Now, the

relations $(2, 3, 7; 4)$ are sufficient[5] in themselves to define the simple group of order 168. Since $Q^2P^5 = Q^{-2}P^5 = (P^2Q^2)^{-1}$, it follows from (5) that the relations $(2, 3, 7; 4)$ imply $(Q^2P^5)^3 = 1$. Therefore, G_{168} is a factor group whenever p is a multiple of four; it satisfies but does not fulfil the conditions $(2, 3, 7; 4p')$; $(Q^2P^5)^{3t'} = 1$.

Suppose now that the period of Q^2P^5 is t (> 3). Then

$$(Q^2P^5)^2 = P^4(P^5Q^2)^{-2},$$

$$(P^2Q^{-2})^{t-2} = P^2Q^2(P^5Q^2)^{t-3},$$

$$(P^2Q^{-2})^{t-4} = Q^4(P^5Q^2)^{t-4}P^5Q^4P^5,$$

$$[Q^4P^5(Q^2P^5)^{t-4}]^2 = 1. \tag{11}$$

Also

$$\begin{aligned}(Q^2P^5)^3 &= Q^2P^5\cdot Q^2P^2\cdot P^3Q^2P^4\cdot P\\ &= Q^2P^3Q^{-2}P^5Q^{-5}\cdot P^4\cdot Q^{-2}P\\ &= Q^2P^3Q^{-2}P^5Q^{-4}P^5\cdot P^5Q^{-1}P\\ &= Q\cdot QP^3\cdot Q^{-2}P^5Q^{-4}P^5\cdot QP^6\cdot Q^{-1}\\ &= QP^4Q^{-3}P^5Q^{-4}P^2Q^{-1}P^3Q^{-1},\end{aligned}$$

$$(Q^2P^5)^3 = (QP^4Q)\cdot Q^{-4}P^5Q^{-4}P^2\cdot (QP^4Q)^{-1}, \tag{12}$$

so that $(Q^2P^5)^3$ has the same period as $Q^4P^5Q^4P^2$.

These last two relations make it possible to show that t may not be four. For then $(Q^4P^5)^2 = 1$; $Q^4P^5 = P^2Q^{-4}$. Substituting in $(Q^4P^5Q^4P^2)^4 = 1$, $P^2 = 1$, which is impossible.

As a corollary to the above results, we shall show that the relations which Coxeter has designated $G^{3,7,7}$ and $G^{3,7,11}$ are incompatible. Coxeter[6] has studied the general defining relations

$$G^{m,n,p}\colon A^m = B^n = C^p = (AB)^2 = (BC)^2 = (CA)^2 = (ABC)^2 = 1$$

and finds that $G^{3,7,p}$ (p odd) may be written $(2, 3, 7; p)$; $(Q^{\frac{1}{2}(p+3)}P^2)^2 = 1$. When $p = 7$, the last relation implies $(Q^2P^5)^2 = 1$, and we have immediate incompatibility. When $p = 11$, the last relation is equivalent to $(Q^4P^5)^2 = 1$. Substituting this result in (12)

$$\begin{aligned}(Q^2P^5)^3 &= QP^4Q^{-3}P^5\cdot Q^{-4}P^2\cdot Q^{-1}P^3Q^{-1}\\ &= QP^4Q^{-3}\cdot P^3\cdot Q^3P^3Q^{-1}\end{aligned}$$

---

[5] Cf. footnote 1.

[6] These results which have been communicated to me by Dr. Coxeter have not yet been submitted for publication.

whence $(Q^2P^5)^3$ is of the same period as P^3. Therefore $(Q^2P^5)^{21} = 1$. From (11)

$$Q^4P^5(Q^2P^5)^{-4}Q^4P^5(Q^2P^5)^{-4} = Q^4P^5Q^4P^5$$

$$P^5(Q^2P^5)^{-5}Q^4P^5(Q^2P^5)^{-5} = (P^5Q^2)^{-4}(Q^2P^5)^{-4} = 1$$

$$(Q^2P^5)^{20} = (P^5Q^2)^{-20},\ P^2Q^{-2} = P^5Q^2,\ Q^4 = P^4$$

which is impossible.

Should Q^2P^5 be of period 5, it follows from (9) that Q^2 is of the same period as P^2, i.e. $Q^{14} = 1$. This condition must be fulfilled, since it is not possible for $Q^2 = 1$ or for $Q^7 = 1$. The latter statement follows immediately from the fact that (2, 3, 7; 7) is of order 1092 and cannot involve any operators of period 5. Hence if $(Q^2P^5)^5 = 1$, Q is of period 14. That these relations are compatible and that they define a group whose order is at least 12,180 is proved by the following generators of $LF(2, 29)$.[7]

$$S = (1, 30, 2)(3, 16, 29)(4, 21, 15)(5, 9, 20)(6, 25, 8)(7, 26, 24)(10, 18, 19)(11, 28, 17)(12, 23, 27)(13, 14, 22),$$

$$T = (1, 28)(2, 27)(3, 26)(4, 25)(5, 24)(6, 23)(7, 22)(8, 21)(9, 20)(10, 19)(11, 18)(12, 17)(13, 16)(14, 15).$$

On the other hand, by using a special enumeration process due to Coxeter and Todd,[8] the author has been able to show that the abstract relations (2, 3, 7); $(Q^2P^5)^5 = 1$ define a group whose order cannot exceed 12,180. Hence *the simple group $LF(2, 29)$ is completely defined by the relations* (2, 3, 7); $t = 5$.

In carrying out the enumeration process, the author found it convenient to use the group $\{Q, PQP^2\}$ of order 28 as the basic subgroup, although it would have been a bit shorter to utilize the icosahedral subgroup $\{QP^2, P^3Q\}$. It is worthy of note that the above definition for $LF(2, 29)$ is considerably simpler than the best existing definition for this group, namely Todd's[9]

$$S^{29} = R^{14} = U^3 = (US)^2 = (UR)^2 = 1, \qquad RS = S^4R$$

4. **A perfect group satisfying** (2, 3, 7; 8). When $t > 5$ and $p > 7$, these two parameters are apparently independent. Hence the simplest case to be considered is $p = 8$, $t = 6$. In this article, we shall determine the group G defined by (2, 3, 7; 8); $(Q^2P^5)^6 = 1$. G will be perfect[10] since the relations (2, 3, 7) alone are sufficient to insure that fact.

It has already been shown that G yields G_{168} as a quotient group. The in-

[7] As linear fractional transformations S and T are:

$$S \equiv \begin{pmatrix} 2, & -3 \\ 1, & -1 \end{pmatrix}, \qquad T \equiv \begin{pmatrix} -1, & 0 \\ 0, & 1 \end{pmatrix}.$$

[8] Proc. Edinburgh Math. Soc. (2), vol. 5 (1937), pp. 26–34.

[9] Journal London Math. Soc., 7 (1932), p. 195.

[10] "Certain Perfect Groups," p. 347.

variant subgroup which gives rise to G_{168} as a factor group is generated by the complete set of conjugates involving Q^4. We study first the operators $P^{-\alpha}Q^4P^{\alpha}$ $(\alpha = 0, 1, 2, \cdots, 6)$.

Since $Q^4P^2Q^4P^5$ is of period two, Q^4 is commutative with $P^2Q^4P^5$ and $P^5Q^4P^2$. Furthermore, $(Q^4P^2Q^4P^5)^2 = 1$ implies $(Q^3P^6Q^3P^5)^2 = 1$. For[11]

$$\begin{aligned} Q^4P^2Q^4P^5 &= Q^3 \cdot QP^2Q^4P \cdot P^4 && (2) \\ &= Q^3P^6Q^{-4}P^5 \cdot Q^{-1}P^4 = Q^3P^6Q^{-4}PQ; \\ Q^3P^6Q^{-4}PQ &= (Q^4P^2Q^4P^5)^{-1} = P^2Q^{-4}P^5Q^{-4} \\ Q^3P^6Q^{-4} \cdot PQ^5P^2Q \cdot Q^3P^5 &= 1, && (2) \end{aligned}$$

$$(Q^3P^6Q^3P^5)^2 = 1. \tag{13}$$

With the aid of this last relation, it can be proved that all the operators $P^{-\alpha}Q^4P^{\alpha}$ are permutable.

$$\begin{aligned} P^5Q^4P^2Q &= P^4 \cdot PQ^4P^2Q = P^4Q^{-1} \cdot P^5Q^4P^6 && (2) \\ &= QPQ^4P^6, \\ Q^4PQ^4P^6 &= Q^3P^5Q^4P^2Q = P \cdot P^6Q^3P^5Q^3 \cdot QP^2Q && (13), (1) \\ &= PQ^5P^2 \cdot Q^5P^2QP = PQ^5 \cdot PQ^7P^5 \cdot Q^3 && (2), (1) \\ &= PQ^4P^{-1}Q^4, \\ (Q^4PQ^4P^6)^2 &= 1. \end{aligned}$$

It follows similarly that $(Q^4P^3Q^4P^4)^2 = 1$. Q^4 is thus seen to be commutative with each of its conjugates under powers of P. Further,

$$\begin{aligned} P^{-\alpha}Q^4P^{\alpha} \cdot P^{-\beta}Q^4P^{\beta} &= P^{-\beta} \cdot P^{\beta-\alpha}Q^4P^{-(\beta-\alpha)}Q^4 \cdot P^{\beta} \\ &= P^{-\beta}Q^4P^{\beta} \cdot P^{-\alpha}Q^4P^{\alpha}, \end{aligned}$$

so that any two operators of the form $P^{-i}Q^4P^i$ are commutative. Hence the group H generated by these seven operators is abelian and involves operators of period two only.

The subgroup H is invariant in G. To prove this fact it is sufficient to show that H is invariant under Q.

$$\begin{aligned} Q^{-1}(P^6Q^4P)Q &= Q^{-1}P^{-3} \cdot P^2Q^4PQ && (2) \\ &= P^3Q \cdot Q^7P^6Q^4P^5 = P^2Q^4P^5, \\ Q^{-1}(P^5Q^4P^2)Q &= Q^{-1}P^{-3} \cdot PQ^4P^2Q && (2) \\ &= P^3Q \cdot Q^7P^5Q^4P^6 = PQ^4P^6. \end{aligned}$$

[11] To facilitate the verification of the manipulation which follows, a number has occasionally been written to the right to indicate the particular relation which is being used in the process of simplification.

Similarly

$$Q^{-1}(P^4Q^4P^3)Q = P^3Q^4P^4,$$

$$Q^{-1}(P^3Q^4P^4)Q = P^4Q^4P^3 \cdot Q^4,$$

$$Q^{-1}(P^2Q^4P^5)Q = P^3Q^4P^4 \cdot P^5Q^4P^2 \cdot Q^4,$$

$$Q^{-1}(PQ^4P^6)Q = P^6Q^4P \cdot P^3Q^4P^4.$$

Therefore, the invariant subgroup of G which is generated by the complete set of conjugates involving Q^4 is abelian and of type $(1, 1, 1, \cdots)$.

The continued product of the seven generators $\prod_{i=0}^{6} P^{2i}Q^4P^{-2i} = (Q^4P^2)^7$. We proceed to show that $(Q^4P^2)^7 = 1$. This will be accomplished by proving that

$$Q^4P^2 = (QP^5Q^6)P(QP^5Q^6)^{-1}, \text{ i.e.}$$

$$Q^3P^2QP = P^5Q^6PQ^2P^3, \tag{2}$$

$$PQ^7P^5 \cdot Q^5P^4 \cdot Q^6P^6Q^2 = 1, \tag{6}$$

$$PQ^7P \cdot Q^3 \cdot P^3Q \cdot P^6Q^2 = 1, \tag{3}$$

$$Q^7 \cdot P^3QP^3 \cdot Q^2 = Q^{16} = 1.$$

It follows then that at most six of the generators of H are independent, and that the order of G cannot exceed $2^6 \cdot 168 = 10{,}752$. That its order is exactly 10,752 is demonstrated by the two generating permutations given below, which yield a transitive representation of G of degree 42.

$$P = (1, 2, 3, 4, 5, 6, 7)(8, 9, 10, 11, 12, 13, 14) \cdots (36, 37, 38, 39, 40, 41, 42),$$

$$Q = (1, 15, 22, 36, 40, 34, 11, 4)(3, 7, 8, 29, 42, 39, 26, 17)(5, 6, 21, 18)$$
$$(9, 24, 10, 16)(12, 20, 23, 32)(13, 28, 35, 27)(14, 31, 25, 19)(30, 37, 38, 33).$$

Therefore, the most general group determined by the relations $(2, 3, 7; 8)$; $(Q^2P^5)^6 = 1$ *is of order* 10,752.

5. **The possible quotient groups of** G. Should any smaller groups than G exist satisfying the same relations, they would be obtainable as quotient groups of G. The invariant subgroups to be considered must be contained in H and will be definable by some additional restrction on P and Q. We consider the possibility of the existence of such groups.

They will arise if it is possible to generate H by fewer than six of the operators $P^{-i}Q^4P^i$. This implies that some product of less than seven of the generators reduces to the identity. Since the product of all seven is the identity, it is sufficient to consider only those products which consist of less than four factors.

Certainly no single operator $P^{-i}Q^4P^i$ can reduce to the identity since Q is of period eight. If

$$P^{-\alpha}Q^4P^{\alpha}\cdot P^{-\beta}Q^4P^{\beta} = 1,$$
$$P^{-(\alpha-\beta)}Q^4P^{(\alpha-\beta)} = Q^4.$$

Q^4 is thus seen to be invariant under P, and the seven generators of H are identical. This contradicts the statement that $\prod_{i=0}^{6} P^{-i}Q^4P^i = 1$.

Suppose then that $P^{-\alpha}Q^4P^{\alpha}\cdot P^{-\beta}Q^4P^{\beta}\cdot P^{-\gamma}Q^4P^{\gamma} = 1$. For purposes of abbreviation, we shall designate $P^{-i}Q^4P^i$ by i. Then since P transforms i into $i + 1$, each of the following sets of three multiplied together yields the identity.

$$\begin{array}{ccc} i & j & k \\ i+1 & j+1 & k+1 \\ \cdots & \cdots & \cdots \\ i+6 & j+6 & k+6 \end{array}$$

By using the fact that no product of fewer than seven of the $P^{-i}Q^4P^i$ reduces to the identity unless it involves either three or four factors, the following properties result:

1. Every triple of the above array has one and only one element in common with every other triple.
2. Every possible pair of numbers is contained in some one of the above triples.
3. No triple not contained in this array may reduce to the identity.

Properties 1 and 2 indicate that i, j, k must correspond to a perfect partition of 7 (in the sense used by Kirkman).[12] Hence the two possible solutions for i, j, k are 1, 2, 4 and 1, 2, 6. It follows from 3 that the invariant subgroup in question (H′) is of order eight, and if we use the solution 1, 2, 4, H′ is composed of the identity and the seven operators

$$P^{-i}(P^{-1}Q^4P\cdot P^{-2}Q^4P^2\cdot P^{-4}Q^4P^4)P^i \qquad (i = 0, 1, 2, \cdots, 6).$$

On equating this expression to the identity, we get the additional restriction which defines H' viz, $Q^4PQ^4P^4Q^4P^2 = 1$, which may be used to define a second group G', fulfilling the conditions (2, 3, 7; 8); $(Q^2P^5)^6 = 1$. This second group is of order 1344, and has been previously obtained in a different connection. In his determination of the transitive groups representable on 14 letters, G. A. Miller[13] found three groups of order 1344. The last of these is isomorphic with G', and the permutations given below were obtained from the representation given by Miller.

$$P = aimegck\cdot bjnfhdl,$$
$$Q = ahcebgdf\cdot ij\cdot km\cdot ln.$$

[12] Transactions Historical Society of Lancashire and Cheshire, vol. 9 (1856–7), p. 128.
[13] Collected Works, vol. 1, p. 215.

It is interesting to note that the list of transitive groups of degree 14 includes a group of order 10,752 which has a number of properties in common with G. However, it is possible to prove that the two groups are abstractly distinct, and it results from the proof that the G^{14}_{10752} cannot be generated in any manner whatever by two operators of periods two and three. Similarly G^{8}_{1344} and the second of the groups G^{14}_{1344} have a number of properties in common with G', but neither of them can be generated by two operators of periods two and three.

It is not difficult to show that the relation $(Q^2P^5)^6 = 1$ is redundant in the definition of G'. If we write $Q^4PQ^4P^4Q^4P^2 = 1$ in the form $Q^4P^5Q^4P^2 \cdot PQ^4P^6 = 1$ we see that $(Q^4P^5Q^4P^2)^2 = 1$. This implies, by the use of (12), that $(Q^2P^5)^6 = 1$.

Had we used the partition 1, 2, 6 instead of 1, 2, 4, the additional restriction would have taken the form $Q^4PQ^4P^2Q^4P^4 = 1$, which is not satisfied by the permutations given for G'. This is as it should be, for it can be shown abstractly that the two definitions for H' are incompatible. However, the two groups of order 1344 are abstractly identical, and the generating permutations given below yield the same group of order 1344 as was used in the foregoing.

$$P = aceikgm \cdot bdfjlhn,$$

$$Q = ak \cdot bl \cdot gh \cdot cfmjdeni.$$

We have then the following theorem: *There are only two groups fulfilling the relations* (2, 3, 7; 8); $(Q^2P^5)^6 = 1$. *The larger, of order* 10,752, *is completely defined by the above relations. The smaller, of order* 1344, *is completely defined by adjoining to the relations* (2, 3, 7; 8) *either* $Q^4PQ^4P^4Q^4P^2 = 1$, *or* $Q^4PQ^4P^2Q^4P^4 = 1$. *In either case,* $(Q^2P^5)^6 = 1$ *is implied as a consequence.*

WASHINGTON, D. C.

ANNALS OF MATHEMATICS
Vol. 38, No. 3, July, 1937

ON EPSTEIN'S ZETA FUNCTION

BY MAX F. DEURING

(Received May 6, 1936)

1. It is the purpose of this note to present some new formulae which connect the zeta functions of quadratic forms or Epstein's zeta-functions to the Riemann zeta-function and functions related to it. Although these formulae did not yield any new results concerning the deeper questions on zeta-functions, I nevertheless believe them of sufficient interest and hope they will be of use in future investigations.

Let $Q(x, y) = ax^2 + bxy + cy^2$ be a quadratic form with real coefficients. We assume Q to be positive definite, therefore

$$a > 0 \qquad \text{and} \qquad D = 4ac - b^2 > 0.$$

The series

$$\tfrac{1}{2} \sum_{n,\, m=-\infty}^{+\infty}\sum{}' Q(n, m)^{-s} = \zeta_Q(s) \tag{1}$$

(the accent indicates the omitting of the term $n = m = 0$) is known to converge absolutely and uniformly for all s in any half plane $\mathrm{R}s = \sigma > 1 + \epsilon$, where $\epsilon > 0$. The function $\zeta_Q(s)$, *Epstein's zeta-function* consequently is regular in the half plane $\sigma > 1$. $\zeta_Q(s)$ is in fact analytic in the whole s-plane, a single pole of order one at $s = 1$ excepted. The functional equation

$$(\tfrac{1}{4}D)^{\frac{1}{2}s}\pi^{-s}\Gamma(s)\,\zeta_Q(s) = (\tfrac{1}{4}D)^{\frac{1}{2}(1-s)}\pi^{-(1-s)}\Gamma(1-s)\zeta_Q(1-s) \tag{2}$$

holds. For these theorems different proofs are known. We shall start our investigations by the development of a formula for $\zeta_Q(s)$ that holds throughout the whole s-plane, giving the analytic continuation of $\zeta_Q(s)$ and at the same time the validity of (2). The same formula will then be used to prove the analytic continuability and the functional equation of the Riemann zeta-function $\zeta(s) = \sum_{n=1}^{\infty} n^{-s}$. Finally we shall use our formula to prove that $\zeta(s)$ vanishes infinitely often on the "critical line" $\sigma = \frac{1}{2}$. This is a well known theorem of Hardy. Our new proof seems to be of some interest because it establishes a relation between the real zeros of $\zeta(\frac{1}{2} + it)$ and the behaviour of $\zeta_Q(\frac{1}{2} + it)$ as $t \to \infty$.

2. We apply the Euler-Poisson summation formula

$$\sum_{n=a}^{b} f(n) = \int_a^b f(x)\,dx + \sum_{\nu=1}^{\infty}\int_a^b f(x)(e^{2\pi i\nu x} + e^{-2\pi i\nu x})\,dx,$$

which is certainly valid if $f(x)$ has a continuous second derivative, to $f(x) = Q(n, m)^{-s}$ for $m \neq 0$

$$\sum_{n=-N}^{+N} Q(n, m)^{-s} = \int_{-N}^{+N} Q(x, m)^{-s}\,dx + \sum_{\nu=1}^{\infty} \int_{-N}^{+N} (e^{2\pi i\nu x} + e^{-2\pi i\nu x})Q(x, m)^{-s}\,dx.$$

Let N tend to $+\infty$, then

$$\sum_{n=-\infty}^{+\infty} Q(n, m)^{-s} = \int_{-\infty}^{+\infty} Q(x, m)^{-s}\,dx + \sum_{\nu=1}^{\infty} \int_{-\infty}^{+\infty} (e^{2\pi i\nu x} + e^{-2\pi i\nu x})Q(x, m)^{-s}\,dx,$$

the convergence of $\sum_{\nu=1}^{M} \int_{-N}^{+N}$ towards $\sum_{\nu=1}^{\infty} \int_{-N}^{+N}$ being uniform with respect to N. The first integral may be evaluated

$$\int_{-\infty}^{+\infty} Q(x, m)^{-s}\,dx = \pi^{\frac{1}{2}} a^{s-1} \left(\frac{D}{4}\right)^{\frac{1}{2}-s} |\, m \,|^{1-2s} \frac{\Gamma(s - \frac{1}{2})}{\Gamma(s)}.$$

Consequently

$$\sum_{n=-\infty}^{+\infty} Q(n, m)^{-s} = \pi^{\frac{1}{2}} a^{s-1} \left(\frac{D}{4}\right)^{\frac{1}{2}-s} |\, m \,|^{1-2s} \frac{\Gamma(s - \frac{1}{2})}{\Gamma(s)} + \sum_{\nu=1}^{\infty} [\omega(s, \nu, m) + \omega(s, -\nu, m)] \tag{3}$$

where $\omega(s, \nu, m)$ is defined by

$$\omega(s, \nu, m) = \int_{-\infty}^{+\infty} e^{2\pi i\nu x} Q(x, m)^{-s}\,dx, \qquad \sigma > 0. \tag{4}$$

3. By an application of Euler's integral

$$\Gamma(s) = \int_0^{\infty} x^{s-1} e^{-x}\,dx$$

we find

$$\Gamma(s)\omega(s, \nu, m) = \int_{-\infty}^{+\infty} e^{2\pi i\nu x}\Gamma(s)Q(x, m)^{-s}\,dx$$
$$= \int_{-\infty}^{+\infty} e^{2\pi i\nu x} \int_0^{\infty} y^{s-1} e^{-yQ(x,m)}\,dy dx = \int_0^{\infty} y^{s-1} \int_{-\infty}^{+\infty} e^{2\pi i\nu x - yQ(x,m)}\,dx dy.$$

For the second integral we have

$$\int_{-\infty}^{+\infty} e^{2\pi i\nu x - y(ax^2+bmx+cm^2)}\,dx$$
$$= e^{-\frac{\pi i bm\nu}{a} - m^2 y \frac{D}{4a} - \frac{\pi^2\nu^2}{ay}} \int_{-\infty}^{+\infty} e^{-ya\left[x - \frac{2\pi i\nu - bmy}{2ay}\right]^2}\,dx$$
$$= \sqrt{\left(\frac{\pi}{ay}\right)}\, e^{-\frac{\pi i bm\nu}{a} - m^2 y \frac{D}{4a} - \frac{\pi^2\nu^2}{ay}}.$$

Hence

$$\Gamma(s)\omega(s, \nu, m) = \sqrt{\left(\frac{\pi}{a}\right)}\, e^{-\frac{\pi i b m \nu}{a}} \int_0^\infty y^{s-\frac{3}{2}} e^{-\frac{m^2 D}{4a} y - \frac{\pi^2 \nu^2}{a y}}\, dy. \tag{5}$$

This representation of $\omega(s, \nu, m)$ shows that $\omega(s, \nu, m)$ is an integral function of s, since the integral in (5) is uniformly convergent in any bounded s-region. Even $\Gamma(s)\,\omega(s, \nu, m)$ is integral, so that $\omega(s, \nu, m) = 0$ for $s = 0, -1, -2, \cdots$, on account of the poles of $\Gamma(s)$. A third expression of $\omega(s, \nu, m)$ is derived from (5) by an application of Mellin's formula

$$\frac{1}{2\pi i} \int_{\alpha - i\infty}^{\alpha + i\infty} \Gamma(u) c^{-u}\, du = e^{-c},$$

valid for $\alpha > 0$ and $\mathrm{R}c > 0$. This formula gives

$$e^{-\frac{\pi^2 \nu^2}{a y}} = \frac{1}{4\pi i} \int_{\alpha - i\infty}^{\alpha + i\infty} \Gamma\left(\frac{u}{2}\right) \left(\frac{\pi^2 \nu^2}{a y}\right)^{-\frac{1}{2}u} du, \qquad \alpha > 0,$$

so that

$$\Gamma(s)\omega(s, \nu, m) = \sqrt{\left(\frac{\pi}{a}\right)}\, e^{-\frac{\pi i b m \nu}{a}} \int_0^\infty y^{s-\frac{3}{2}} e^{-\frac{m^2 D y}{4a}} \frac{1}{4\pi i} \int_{\alpha - i\infty}^{\alpha + i\infty} \Gamma\left(\frac{u}{2}\right) \left(\frac{\pi^2 \nu^2}{a y}\right)^{-\frac{1}{2}u} du\, dy.$$

Provided that $\sigma - \frac{1}{2} + \frac{1}{2}\alpha > 0$, we may change the order of integration for then the inner integral $\int_0^\infty \cdots dy$ is convergent. Using once more Euler's integral, we obtain

$$\begin{aligned} &(\tfrac{1}{4}D)^{\frac{1}{2}s} \pi^{-s} \Gamma(s)\omega(s, \nu, m) \\ &= \frac{\pi^{-\frac{1}{2}-s}}{4i} \left(\frac{D}{4}\right)^{\frac{1-s}{2}} a^{s-1} e^{-\frac{\pi i b m \nu}{a}} \int_{\alpha - i\infty}^{\alpha + i\infty} \left(\frac{D\pi^2}{4a^2}\right)^{-\frac{1}{2}u} \Gamma\left(\frac{u}{2}\right) \Gamma\left(\frac{u}{2} + s - \frac{1}{2}\right) \\ &\qquad\qquad | \nu |^{-u} | m |^{-u-2s+1}\, du. \end{aligned} \tag{6}$$

The two conditions $\sigma > 0$ and $\sigma - \frac{1}{2} + \frac{1}{2}\alpha > 0$ mean that the poles of the integrand are on the left of the line of integration. Introducing a new variable v by $u + 2s - 1 = v$, we get

$$\begin{aligned} &(\tfrac{1}{4}D)^{\frac{1}{2}s} \pi^{-s} \Gamma(s)\omega(s, \nu, m) \\ &= \frac{\pi^{-\frac{1}{2}-(1-s)}}{4i} \left(\frac{D}{4}\right)^{\frac{s}{2}} a^{-s} e^{-\frac{\pi i b m \nu}{a}} \int_{\alpha+2\sigma-1-i\infty}^{\alpha+2\sigma-1+i\infty} \left(\frac{D\pi^2}{4a^2}\right)^{-\frac{1}{2}v} \Gamma\left(\frac{v}{2} + (1-s) - \frac{1}{2}\right) \\ &\qquad\qquad \Gamma\left(\frac{v}{2}\right) | \nu |^{-v-2(1-s)+1} | m |^{-v}\, dv \end{aligned}$$

where again the line of integration is on the right of the poles of the integrand. Comparing this result to (6), we obtain

$$(\tfrac{1}{4}D)^{\frac{1}{2}s} \pi^{-s} \Gamma(s)\omega(s, \nu, m) = (\tfrac{1}{4}D)^{\frac{1}{2}(1-s)} \pi^{-(1-s)} \Gamma(1-s)\omega(1-s, m, \nu). \tag{7}$$

4. (6) shows that in any bounded s-region an inequality

$$| \omega(s, \nu, m) | < \text{const.} \, | \nu |^{-\alpha} \, | m |^{1-\alpha-2\sigma}$$

holds uniformly for all ν and m. The series

$$\tag{8} \tfrac{1}{2} \sum_{m=-\infty}^{+\infty}{}' \sum_{\nu=1}^{\infty} (\omega(s, \nu, m) + \omega(s, -\nu, m)) = R_Q(s)$$

is therefore absolutely and uniformly convergent in any bounded s-region and represents an integral function $R_Q(s)$. From (2) we have

$$\tag{9} \zeta_Q(s) = a^{-s}\zeta(2s) + a^{s-1}\sqrt{\pi}\left(\frac{D}{4}\right)^{\frac{1}{2}-s} \frac{\Gamma(s-\frac{1}{2})}{\Gamma(s)} \zeta(2s-1) + R_Q(s).$$

This, together with (6) and (8) is the formula announced in 1. If we assume the meromorphic character of the Riemann zeta-function $\zeta(s)$, we derive from (9) that $\zeta_Q(s)$ is analytic in the whole s-plane with only one pole of order one at $s = 1$, which in (9) is given by the pole of $\zeta(2s - 1)$. Form (7) and (8) the functional equation

$$\tag{10} (\tfrac{1}{4}D)^{\frac{1}{2}s}\pi^{-s}\Gamma(s)R_Q(s) = (\tfrac{1}{4}D)^{\frac{1}{2}(1-s)}\pi^{-(1-s)}\Gamma(1-s)R_Q(1-s)$$

follows. Using (9) and expressing $\zeta(2s - 1)$ through $\zeta(2 - 2s)$ by means of Riemann's equation, we find (2).

5. It is interesting to note, that the two properties of $\zeta(s)$ used in this argument—the meromorphic character and the functional equation—may be derived too from our formulae. Consider the special function $\zeta_Q(s)$ belonging to $Q(x, y) = x^2 + \xi^2 y^2$. For simplicity write $\zeta(s, \xi)$ instead of $\zeta_Q(s)$, similarly $R(s, \xi)$ for $R_Q(s)$. We have $a = 1$ and $D = 4\xi^2$. The definition of $\zeta(s, \xi)$ gives

$$\zeta(s, \xi) = \xi^{-2s}\zeta\left(s, \frac{1}{\xi}\right).$$

In this equation we express $\zeta(s, \xi)$ and $\zeta(s, 1/\xi)$ by means of (9) and find thus

$$\tag{11} R\left(s, \frac{1}{\xi}\right) = \xi^{2s}R(s, \xi) + (\xi^{2s} - 1)\zeta(2s) + (\xi - \xi^{2s-1})\sqrt{\pi}\,\frac{\Gamma(s-\frac{1}{2})}{\Gamma(s)}\zeta(2s-1).$$

Since we do not assume the meromorphic character of $\zeta(s)$, this is proved only for $\sigma > 1$. Put $2s = w$, then

$$\zeta(w-1) = \frac{\Gamma(\frac{1}{2}w)\cdot\pi^{-\frac{1}{2}}}{(\xi - \xi^{w-1})\Gamma\left(\frac{w-1}{2}\right)}\left[R\left(\frac{w}{2}, \frac{1}{\xi}\right) - \xi^{w}R\left(\frac{w}{2}, \xi\right) + (1 - \xi^{w})\zeta(w)\right].$$

This formula, for arbitrary $\xi \neq 1$, evidently supplies the continuation of $\zeta(w)$ into $Rw > \alpha - 1$, if $\zeta(w)$ is known to be analytic in $Rw > \alpha$. Starting with $\alpha = 2$, we may by steps find out the meromorphic character of $\zeta(w)$.

6. We substitute the value of $\omega(s, \nu, m)$ given by (6) into

$$R(s, \xi) = \tfrac{1}{2} \sum_{m=-\infty}^{+\infty}{}' \sum_{\nu=-\infty}^{+\infty} [\omega(s, \nu, m) + \omega(s, -\nu, m)].$$

Under the restrictions $\sigma > 1$, $\alpha + 2\sigma - 1 > 1$ the summations may be taken under the integral so that

$$\xi^{2s-1}\Gamma(s)R(s, \xi) = \frac{1}{2i\sqrt{\pi}} \int_{\alpha-i\infty}^{\alpha+i\infty} (\pi\xi)^{-u}\, \Gamma\left(\frac{u}{2}\right)\Gamma\left(\frac{u}{2} + s - \frac{1}{2}\right)\zeta(u)\zeta(u + 2s - 1)\, du.$$

Put

$$\pi^{-\frac{s}{2}}\Gamma\left(\frac{s}{2}\right)\zeta(s) = \varphi(s).$$

Then

$$\xi^{2s-1}\pi^{-s}\Gamma(s)R(s, \xi) = \frac{1}{2\pi i} \int_{\alpha-i\infty}^{\alpha+i\infty} \xi^{-u}\varphi(u)\varphi(u + 2s - 1)\, du. \tag{12}$$

We now consider the inverse of (12) in the sense of Mellin[1])

$$\frac{\pi^s}{\Gamma(s)}\varphi(u)\varphi(u + 2s - 1) = \int_0^\infty \xi^{u+2s-2} R(s, \xi)\, d\xi. \tag{13}$$

This equation may be used to prove Riemann's functional equation

$$\varphi(u) = \varphi(1 - u)$$

in a manner similar to that adopted by Riemann himself in his proof based upon

$$\pi^{-\frac{s}{2}}\Gamma\left(\frac{s}{2}\right)\zeta(s) = \int_0^\infty \xi^{\frac{s}{2}-1} \sum_{n=1}^{\infty} e^{-\pi n^2 \xi}\, d\xi.$$

We divide the integral in (13) into $\int_0^1$ and $\int_1^\infty$ and use $\eta = 1/\xi$ as a new variable in $\int_0^1$. Expressing $R(s, \xi) = R(s, 1/\eta)$ through $R(s, \eta)$ by means of (10) we obtain

$$\begin{aligned}
\int_0^1 \xi^{u+2s-2} R(s, \xi)\, d\xi &= \int_1^\infty \eta^{-u-2s} R\left(s, \frac{1}{\eta}\right) d\eta \\
&= \int_1^\infty \eta^{-u} R(s, \eta)\, d\eta + \zeta(2s) \int_1^\infty (\eta^{-u} - \eta^{-u-2s})\, d\eta \\
&\qquad + \sqrt{\pi}\,\frac{\Gamma(s - \frac{1}{2})}{\Gamma(s)}\,\zeta(2s - 1) \int_1^\infty (\eta^{1-u-2s} - \eta^{-u-1})\, d\eta \\
&= \int_1^\infty \eta^{-u} R(s, \eta) d\eta + \zeta(2s)\left[\frac{1}{u - 1} - \frac{1}{u + 2s - 1}\right] \\
&\qquad + \sqrt{\pi}\,\frac{\Gamma(s - \frac{1}{2})}{\Gamma(s)}\left[\frac{1}{u + 2s - 2} - \frac{1}{u}\right]\zeta(2s - 1).
\end{aligned}$$

---

[1] See for instance: S. Bochner, *Vorlesungen über Fouriersche Integrale*, Leipzig 1932, page 148.

We finally express $R(s, \eta)$ through $R(1 - s, \eta)$ according to (10) and introduce the new value of $\int_0^1$ into (13)

$$\varphi(u)\varphi(u + 2s - 1) = \pi^{-s}\Gamma(s) \int_1^\infty \xi^{u+2s-2} R(s, \xi)\, d\xi$$

$$(14) \qquad + \pi^{-(1-s)}\Gamma(1 - s) \int_1^\infty \xi^{-u-2s} R(1 - s, \xi) d\xi$$

$$+ \varphi(2s)\left[\frac{1}{u - 1} - \frac{1}{u + 2s - 1}\right] + \varphi(2s - 1)\left[\frac{1}{u + 2s - 2} - \frac{1}{u}\right].$$

In any region $| s | + | u | <$ const. both integrals are uniformly convergent with respect to both variables s and u, since from (12) $| R(s, \xi) | <$ const. $\xi^{-\alpha}$ follows for arbitrary $\alpha > 0$. Both integrals therefore are integral functions of s and u. (14) has been proved under the assumptions $\sigma > 1$, $\mathrm{R}u + 2\sigma - 1 > 1$ and the existence of $\zeta(s)$ has been used only for $\sigma > 1$. Consequently (14) yields again the analytic continuation of $\zeta(s)$: In order to continue $\varphi(u)$ into $\mathrm{R}u > -\delta$, choose s such that $-\delta + 2\sigma - 1 > 1, \sigma > 1$, then (14) expresses $\varphi(u)$ in $\mathrm{R}u > -\delta$ by means of known functions.. We shall take $s = \frac{1}{2}$ in (14). In order to do this we need the connection between $\zeta(0)$ and the residue ρ of $\zeta(s)$ at $s = 1$. Take $s = 0$ in (11). Since $R(0, \xi) = 0$ by (5) we have

$$(\xi - 1)\left(\frac{\rho}{2s - 1} + \cdots\right) + (\xi - 1)\left(\frac{1}{s - \frac{1}{2}} + \cdots\right)\zeta(0) = 0,$$

$$\zeta(0) = -\tfrac{1}{2}\rho.$$

The expansions of $\varphi(s)$ into powers of $s - 1$ and s respectively therefore begin as follows

$$(15) \qquad \varphi(s) = \frac{\rho}{s - 1} + a + \cdots, \qquad \varphi(s) = -\frac{\rho}{s} + b + \cdots.$$

From (15) we derive the value of $\varphi(2s)\left[\dfrac{1}{u - 1} - \dfrac{1}{u + 2s - 1}\right]$ at $s = \frac{1}{2}$ to be

$$\frac{\rho}{u^2} + a\left[\frac{1}{u - 1} - \frac{1}{u}\right],$$

and similarly the value of $\varphi(2s - 1)\left[\dfrac{1}{u + 2s - 2} - \dfrac{1}{u}\right]$ at $s = \frac{1}{2}$ is

$$\frac{\rho}{(u - 1)^2} + b\left[\frac{1}{u - 1} - \frac{1}{u}\right].$$

Now (14) gives for $s = \frac{1}{2}$

$$(16) \qquad \varphi(u)^2 = \int_1^\infty \xi^{u-1} R(\tfrac{1}{2}, \xi) d\xi + \int_1^\infty \xi^{-u} R(\tfrac{1}{2}, \xi) d\xi + \rho(u^{-2} + (1 - u)^{-2}) - (a + b)(u^{-1} + (1 - u)^{-1}).$$

The expansion of $\varphi(u)^2$ into powers of u begins by (16)

$$\varphi(u)^2 = \rho/u^2 + \cdots,$$

but by (15)

$$\varphi(u)^2 = \rho^2/u^2 + \cdots.$$

Consequently $\rho = 1$, since $\rho \neq 0$. The right hand side of (16) remains unchanged if u is replaced by $1 - u$. Therefore $\varphi(u)^2 = \varphi(1-u)^2$. Whether $\varphi(u) = \varphi(1-u)$ or $\varphi(u) = -\varphi(1-u)$ we decide by taking $s = 1 - u$ in the first of the equations (15) and $s = u$ in the second. Comparison yields $\varphi(u) = \varphi(1-u)$. This is Riemann's functional equation.

7. From (9) and (12) we derive

$$\zeta(s, \xi) = \zeta(2s) + \pi^{\frac{1}{2}} \frac{\Gamma(s - \frac{1}{2})}{\Gamma(s)} \zeta(2s-1)\xi^{1-2s} + \frac{\xi^{1-2s}\pi^{-s}}{2\pi i \Gamma(s)} \int_{\alpha - i\infty}^{\alpha + i\infty} \xi^{-u} \varphi(u)\varphi(u + 2s - 1)\, du.$$

Write $\Phi(s, \xi)$ for $\xi^s \pi^{-s} \Gamma(s) \zeta(s, \xi)$, so that

$$\Phi(s, \xi) = \xi^s \varphi(2s) + \xi^{1-s}\varphi(2s-1) + \xi^{1-s} \frac{1}{2\pi i} \int_{\alpha - i\infty}^{\alpha + i\infty} \xi^{-u}\varphi(u)\varphi(u + 2s - 1)\, du. \tag{17}$$

Let s be a point on the line $\sigma = \frac{1}{2}$. We shift the line of integration to $\mathrm{R}u = \frac{1}{2}$. The two poles $u = 1$ and $u = 2 - 2s$ are passed over, so that

$$\begin{aligned} \Phi(\tfrac{1}{2} + it, \xi) = {} & (\xi^{\frac{1}{2}+it} + \xi^{-\frac{1}{2}-it})\varphi(1 + 2it) + (\xi^{\frac{1}{2}-it} + \xi^{-\frac{1}{2}+it})\varphi(1 - 2it) \\ & + \int_{-\infty}^{+\infty} \xi^{-i(y+t)} \varphi(\tfrac{1}{2} + iy)\varphi(\tfrac{1}{2} + i(y + 2t))\, dy. \end{aligned} \tag{18}$$

$\Phi(\frac{1}{2} + it, \xi)$ and $\varphi(\frac{1}{2} + it)$ are real valued functions of the real variable t. This is an immediate consequence of the functional equations for $\zeta(s, \xi)$ and $\zeta(s)$.

We use (18) to prove Hardy's theorem that $\zeta(\frac{1}{2} + it)$ or else $\varphi(\frac{1}{2} + it)$ has infinitely many real zeros. Suppose on the contrary that $\zeta(\frac{1}{2} + it)$ vanishes only for a finite number of t's. Then for $t > T_0$, $\varphi(\frac{1}{2} + it)$ is of constant sign. Take $t > T_0$. Then the integral in (18) is for the special case $\xi = 1$ equal to

$$\int_{-T_0}^{\infty} \varphi(\tfrac{1}{2} + iy)\varphi(\tfrac{1}{2} + i(y + 2t))\, dy + \int_{-\infty}^{-2t+T_0} \varphi(\tfrac{1}{2} + iy)\varphi(\tfrac{1}{2} + i(y + 2t))\, dy$$
$$+ \int_{-2t+T_0}^{-T_0} |\varphi(\tfrac{1}{2} + iy)\varphi(\tfrac{1}{2} + i(y + 2t))|\, dy.$$

We shall prove that

$$\int_{-2t+T_0}^{-T_0} |\varphi(\tfrac{1}{2} + iy)\varphi(\tfrac{1}{2} + i(y + 2t))|\, dy > \text{const. } e^{-\frac{1}{2}\pi t} t^{\frac{1}{2}}, \tag{19}$$

but that the other two integrals as well as $\varphi(1 + 2it) + \varphi(1 - 2it)$ are $o(e^{-\frac{1}{2}\pi t} t^{\frac{1}{2}})$.

So

$$\Phi(\tfrac{1}{2} + it, 1) > \text{const. } e^{-\frac{1}{2}\pi t} t^{\frac{1}{2}}$$

for sufficiently large t. But this is a contradiction[2] to

$$\Phi(\tfrac{1}{2} + it, \xi) = o(t^{\frac{1}{2}}).$$

It is interesting to note, that (19) contradicts too the existence of infinitely many real zeros of $\Phi(\frac{1}{2} + it, 1)$, which might be proved in a separate way.

In order to prove (19) we use

$$(20) \qquad |\Gamma(\sigma + it)| = b(\sigma)\,|t|^{-\frac{1}{4}} e^{-\frac{1}{2}\pi|t|}(1 + O(|t|^{-1})) \text{ as } |t| \to \infty,$$

a well known consequence of Stirling's asymptotic formula for $\Gamma(s)$. From (20) we derive that

$$|\Gamma(\tfrac{1}{4} + i\tfrac{1}{2}y)| \geqq \text{const. } |y|^{-\frac{1}{4}} e^{-\frac{1}{4}\pi|y|}$$

and

$$|\Gamma(\tfrac{1}{4} + i\,\tfrac{1}{2}(y + 2t))| \geqq \text{const. } |y + 2t|^{-\frac{1}{4}} e^{-\frac{1}{4}\pi|y+2t|}$$

if $-T_0 \geqq y \geqq -2t + T_0$. Therefore

$$\int_{-2t+T_0}^{-T_0} |\varphi(\tfrac{1}{2} + iy)\varphi(\tfrac{1}{2} + i(y + 2t))|\,dy$$

$$\geqq \text{const. } e^{-\frac{1}{2}\pi t} \int_{-2t+T_0}^{-T_0} |\zeta(\tfrac{1}{2} + it)\zeta(\tfrac{1}{2} + i(y + 2t))|\,|y|^{-\frac{1}{4}}|y + 2t|^{-\frac{1}{4}}\,dy.$$

$|y(y + 2t)|^{-\frac{1}{4}}$ takes its minimum for $y = -t$, so that

$$\int_{-2t+T_0}^{-T_0} |\varphi(\tfrac{1}{2} + iy)\varphi(\tfrac{1}{2} + i(y + 2t))|\,dy$$

$$\geqq \text{const. } e^{-\frac{1}{2}\pi t} t^{-\frac{1}{2}} \int_{-2t+T_0}^{-T_0} |\zeta(\tfrac{1}{2} + iy)\zeta(\tfrac{1}{2} + i(y + 2t))|\,dy$$

$$\geqq \text{const. } e^{-\frac{1}{2}\pi t} t^{-\frac{1}{2}} \left| \int_{-2t+T_0}^{-T_0} \zeta(\tfrac{1}{2} + iy)\zeta(\tfrac{1}{2} + i(y + 2t))\,dy \right|$$

$$= \text{const. } e^{-\frac{1}{2}\pi t} t^{-\frac{1}{2}} \left\{ -i \int_{\frac{1}{2}-i(2t-T_0)}^{2-i(2t-T_0)} \zeta(z)\zeta(z + 2it)\,dz - i \int_{2-i(2t-T_0)}^{2-iT_0} \zeta(z)\zeta(z + 2it)\,dz \right.$$

$$\left. - i \int_{2-iT_0}^{\frac{1}{2}-iT_0} \zeta(z)\zeta(z + 2it)\,dz \right\}.$$

The first integral is $O(t^{\frac{1}{4}})$. For $\zeta(z + 2it)$ runs through a set of values independent of t and $\zeta(z) = O(t^{\frac{1}{4}})$ by a well known theorem. We need in fact only

[2] See for instance: E. C. Titchmarsh, *On Epstein's Zeta Function* Proc. London Math. Soc. II ser. 36 (1934) 485–500.

$\zeta(z) = o(t)$, which is rather easy to prove. In the same way the third integral is shown to be $O(t^{\frac{1}{2}})$. The second integral is equal to

$$\sum_{n,m} n^{-2} m^{-2} \int_{-2t+T_0}^{-T_0} n^{-iy} m^{-i(y+2t)}\, dy$$

$$= 2(t - T_0) + \sum_{n,m}{}' \frac{m^{-2it}}{n^2 m^2} \int_{-2t+T_0}^{-T_0} e^{-iy \log nm}\, dy$$

$$= 2(t - T_0) + \sum_{n,m}{}' \frac{m^{-2it}((nm)^{i(2t-T_0)} - (nm)^{iT_0})}{n^2 m^2 \log nm}$$

$$= 2(t - T_0) + O(1).$$

This proves (19).

From $\zeta(1 + it) = O(\log t)$ ($\zeta(1 + it) = O(t^{\frac{1}{2}})$ would suffice) we derive

$$\varphi(1 + 2it) + \varphi(1 - 2it) = O(e^{-\frac{1}{4}\pi t} \log t),$$

To complete our proof of $\Phi(\frac{1}{2} + it, 1) > \text{const. } e^{-\frac{1}{4}\pi t} t^{\frac{1}{2}}$, we have to show only

$$\int_{-T_0}^{\infty} \varphi(\tfrac{1}{2} + iy)\varphi(\tfrac{1}{2} + i(y + 2t))\, dy + \int_{-\infty}^{-2t+T_0} \varphi(\tfrac{1}{2} + iy)\varphi(\tfrac{1}{2} + i(y + 2t))\, dy = O(e^{-\frac{1}{4}\pi t} t^{-\frac{1}{2}}).$$

Since the two integrals are conjugate imaginaries, we need only to consider the first one. By (20),

$$\int_{-T_0}^{T_0} \varphi(\tfrac{1}{2} + iy)\varphi(\tfrac{1}{2} + i(y + 2t))\, dy = O\left(\int_{-T_0}^{+T_0} \varphi(\tfrac{1}{2} + iy) e^{-\frac{1}{4}\pi |y+2t|} \,|\, y + 2t \,|^{-\frac{1}{2}}\, dy\right)$$
$$= O(e^{-\frac{1}{4}\pi t} t^{-\frac{1}{2}}),$$

and

$$\int_{T_0}^{\infty} \varphi(\tfrac{1}{2} + iy)\varphi(\tfrac{1}{2} + i(y + 2t))\, dy = O\left(\int_{T_0}^{\infty} e^{-\frac{1}{4}\pi y - \frac{1}{4}\pi t} (y(y + 2t))^{-\frac{1}{2}}\, dy\right)$$
$$= O(e^{-\frac{1}{4}\pi t} t^{-\frac{1}{2}}).$$

q.e.d.

JENA, GERMANY.

Annals of Mathematics
Vol. 38, No. 3, July, 1937

ABSOLUTE REGULARITY AND THE NÖRLUND MEAN

By Florence M. Mears

(Received September 14, 1936)

1. **Introduction.** For any given series, $\sum_{k=1}^{\infty} u_k$, with u_k real or complex, form the sequence $\{U_k\}$, where $U_k = \sum_{n=1}^{k} u_n$. We shall consider sequence to sequence transformations of the type

$$U'_n = \sum_{k=1}^{\infty} a_{nk} U_k,$$

in which the elements of the matrix $\| a_{nk} \|$ are real or complex constants. If the series $\sum_{k=1}^{\infty} | u'_k |$ is convergent, where $u'_k = U'_k - U'_{k-1}$, we shall say that $\sum_{k=1}^{\infty} u_k$ is absolutely summable[1] by the transformation. We shall say that the matrix defines an absolutely regular method of summability A, if every absolutely convergent series $\sum_{k=1}^{\infty} u_k$ is transformed into an absolutely convergent series $\sum_{k=1}^{\infty} u'_k$, and that the method A includes absolutely the method B, defined similarly by $\| b_{nk} \|$, if every series summed absolutely by B is summed absolutely by A. Paragraph 2 includes the necessary and sufficient conditions for absolute regularity.

Let $\{a_n\}$ be a sequence of complex numbers, such that $A_k = \sum_{n=1}^{k} a_n \neq 0$. The series $\sum_{k=1}^{\infty} u_k$ is said to be summable to U' by the Nörlund mean A, if

$$\lim_{n\to\infty} U'_n = \lim_{n\to\infty} [A_n^{-1} \textstyle\sum_{k=1}^{n} a_{n-k+1} U_k]$$

exists and is equal to U'.[2] In a recent paper,[3] the writer proved three multiplication theorems for the Nörlund mean. Paragraph 3 presents a fourth multipli-

[1] This is the same as the absolute summability defined for the Cesàro and Riesz means by Fekete, Matematikai és Természettudományi Értesitö, vol. 32 (1914), pp. 389–425, and Obrechkoff, Comptes Rendus, vol. 185 (1928), pp. 215–217, respectively.

[2] Nörlund, Lunds Universitet, Årsskrift, (2), vol. 16 (1920), No. 3.

[3] Bulletin American Mathematical Society, vol. 41 (1935), pp. 875–880. In Theorems 2 and 3 of this paper it is possible to replace the condition for regularity of the Nörlund means A and B by less restrictive ones. In Theorem 2, we substitute the assumption that the transformation D includes A and B; in Theorem 3, the assumption that C includes A and B. For these two theorems, the terms of the Nörlund sequences need not be restricted to non-negative numbers, but may be complex, such that $A_k \neq 0$, $B_k \neq 0$, with the additional restriction that $D_k \neq 0$ for Theorem 2, and $C_k \neq 0$ for Theorem 3. The proof of Theorem 2 remains unchanged. In Theorem 3, the transformation defined by $\| a_{nk} \|$ is regular since C includes A; therefore (a) and (b) of the lemma are satisfied. Condition (c) of the lemma is satisfied since $C = C'B$. With these changes, Theorem 2 includes Chapman's extension of Cesàro's multiplication theorem. (Chapman, Proceedings of the London Mathematical Society (2), vol. 9, 1910, p. 378). It is obvious from Theorem 3,

cation theorem, which includes Cauchy's theorem for the multiplication of absolutely convergent series, and its extension by Kogbetliantz to the Cesàro mean.[4]

2. **Absolute Regularity.** In the following proof, let $a_{0p} = A_{0p} = 0$ for all p.

THEOREM I. *The necessary and sufficient conditions that* $\sum_{n=1}^{\infty} | u'_n |$, *defined by* $U'_n = \sum_{k=1}^{\infty} a_{nk} U_k$, *shall converge whenever* $\sum_{n=1}^{\infty} | u_n |$ *converges are*

(1) $\sum_{k=1}^{\infty} a_{nk}$ *converges, for all* n;

(2) $\sum_{n=1}^{\infty} | \sum_{p=k}^{\infty} (a_{np} - a_{n-1,p}) | \leqq C$, *for all* k, *where* C *is a positive constant.*

Moreover, $\sum_{n=1}^{\infty} | u'_n | \leqq C \sum_{n=1}^{\infty} | u_n |$.

PROOF. The conditions are sufficient. By (1), $\sum_{k=k_1}^{\infty} a_{nk}$ converges for each k_1; we represent it by A_{nk_1}. Then for each n, U'_n is defined, since $U'_n = \sum_{k=1}^{\infty} A_{nk} u_k$, which converges. We denote $A_{nk} - A_{n-1,k}$ by S_{nk}. By (2)

$$\sum_{n=1}^{\infty} | u'_n | = \sum_{n=1}^{\infty} \left| \sum_{k=1}^{\infty} S_{nk} u_k \right| \leqq \sum_{k=1}^{\infty} \sum_{n=1}^{\infty} | S_{nk} u_k | \leqq C \sum_{k=1}^{\infty} | u_k |;$$

therefore $\sum_{n=1}^{\infty} | u'_n |$ converges.

The conditions are necessary. Choosing $U_k = 1$ for all k, we have $U'_n = \sum_{k=1}^{\infty} a_{nk}$; in order that U' be defined for each n, (1) is necessary.

If (2) is not satisfied, then either

(2a) for some k, say k_1, $\sum_{n=1}^{\infty} | S_{nk} |$ diverges, or

(2b) for each k, $\sum_{n=1}^{\infty} | S_{nk} |$ converges, but

$$\limsup_{k \to \infty} \sum_{n=1}^{\infty} | S_{nk} | = \infty.$$

To prove (2a) impossible, we choose $u_n = 0$, $n \neq k_1$; $u_{k_1} = 1$. Then

$$\sum_{n=1}^{\infty} | u'_n | = \sum_{n=1}^{\infty} | S_{nk_1} u_{k_1} |,$$

which diverges.

To prove (2b) impossible, we assume it true, and construct an absolutely convergent series $\sum_{k=1}^{\infty} u_k$, $| u_{k_m} | < 2^{-m}$, $m = 1, 2, \cdots$, and $u_k = 0$, $k \neq k_m$, such that $\sum_{n=1}^{\infty} | u'_n |$ diverges. We have $\sum_{k=1}^{n} | u_k | < 1$ for all n.

Since $\limsup_{k \to \infty} \sum_{n=1}^{\infty} | S_{nk} | = \infty$, we can choose k_1 so that $\sum_{n=1}^{\infty} | S_{nk_1} | > 2$; since by (2a) $\sum_{n=1}^{\infty} | S_{nk_1} |$ converges, we can find n_1, such that $\sum_{n=n_1+1}^{\infty} | S_{nk_1} | < 1$. Then $\sum_{n=1}^{n_1} | S_{nk_1} | > 1$. Since $\{U_n\}$ must be bounded, there exists a

---

that in the following theorem by Belinfante the restriction $r \geqq 1$ may be replaced by $r > 0$: If $\sum_{n=1}^{\infty} u_n$ is summable (C, s) to U, and if $\sum_{n=1}^{\infty} v_n$ is summable (C, r) to V, and bounded $(C, r-1)$, $(s \geqq 0, r \geqq 1)$, the product series $\sum_{n=1}^{\infty} w_n$ is summable $(C, r+s)$ to UV. (Belinfante, Koninklijke Akademie te Amsterdam, Verslag, vol. 32 (1923), pp. 177–189, 523–535).

[4] Kogbetliantz, Mémorial des Sciences Mathématiques, No. 51, p. 27.

constant K such that $|A_{nk}| < K$ for all n and k.[5] It follows that we can choose $M_1 > 2$, so that for all k,

$$\sum_{n=1}^{n_1} |S_{nk}| \leqq 2 \sum_{n=1}^{n_1} |A_{nk}| < 2n_1 K < M_1.$$

We choose $|u_{k_1}| = (M_1)^{-1}$; we shall choose u_{k_m}, $m \geqq 2$, so that $\sum_{k=k_2}^{\infty} |u_k| < (M_1)^{-1}$. We have

$$\begin{aligned}\sum_{n=1}^{n_1} |u_n'| &= \sum_{n=1}^{n_1} \left| \sum_{k=1}^{\infty} S_{nk} u_k \right| \\ &\geqq \sum_{n=1}^{n_1} |S_{nk_1} u_{k_1}| - \sum_{k=k_2}^{\infty} \sum_{n=1}^{n_1} |S_{nk} u_k| \\ &> (M_1)^{-1} - M_1(M_1)^{-1} = (M_1)^{-1} - 1.\end{aligned}$$

Let m be an integer greater than 1. We choose $k_m > k_{m-1}$, so that $\sum_{n=1}^{\infty} |S_{nk_m}| > 2M_{m-1}^m + M_{m-1} + 1$, and $n_m > n_{m-1}$ so that $\sum_{n=n_m+1}^{\infty} |S_{nk_m}| < 1$. We choose $M_m > M_{m-1}$ so that for all k, $\sum_{n=1}^{n_m} |S_{nk}| < M_m$. Then since $\sum_{n=1}^{n_{m-1}} |S_{nk_m}| < M_{m-1}$, we have $\sum_{n_{m-1}+1}^{n_m} |S_{nk_m}| > 2M_{m-1}^m$. We choose $|u_{k_m}| = (M_{m-1})^{-m}$; we shall choose $\sum_{k=k_m+1}^{\infty} |u_k| < (M_m)^{-m}$. We have

$$\begin{aligned}\sum_{n_{m-1}+1}^{n_m} |u_n'| &= \sum_{n_{m-1}+1}^{n_m} \left| \sum_{k=1}^{\infty} S_{nk} u_k \right| \\ &\geqq P - Q - R,\end{aligned}$$

where

$$\begin{aligned}P &= \sum_{n_{m-1}+1}^{n_m} |S_{nk_m} u_{k_m}| > 2, \\ Q &= \sum_{p=1}^{m-1} \sum_{n_{m-1}+1}^{n_m} |S_{nk_p} u_{k_p}| \leqq \sum_{p=1}^{m-1} \sum_{n_p+1}^{\infty} |S_{nk_p} u_{k_p}| < 1, \\ R &= \sum_{p=m+1}^{\infty} \sum_{n_{m-1}+1}^{n_m} |S_{nk_p} u_{k_p}| \leqq \sum_{p=m+1}^{\infty} \sum_{n=1}^{n_m} |S_{nk_p} u_{k_p}| < M_m(M_m)^{-m}.\end{aligned}$$

Hence

$$\sum_{n_{m-1}+1}^{n_m} |u_n'| > 1 - (M_m)^{1-m} > 1 - 2^{1-m},$$

and

$$\begin{aligned}\sum_{n=1}^{n_m} |u_n'| &= \sum_{n=1}^{n_m} \left| \sum_{k=1}^{\infty} S_{nk} u_k \right| \\ &> (M_1)^{-1} - 1 + \sum_{k=2}^{m} (1 - 2^{1-k}) \\ &= (M_1)^{-1} + m - 3 + 2^{1-m}.\end{aligned}$$

It follows that $\sum_{n=1}^{\infty} |u_n'|$ diverges, which completes the theorem.

---

[5] Hahn, Monatshefte für Mathematik und Physik, vol. 32 (1922), p. 29.

If the terms of the sequence defining the Nörlund mean A are real and non-negative, it is possible to state the following simple conditions for absolute regularity.

THEOREM II. *A sufficient condition for the absolute regularity of the Nörlund mean A, where $a_n \geqq 0$, is the existence of a finite constant p such that*

$$(1)\ A_n^2 \geqq A_{n-1}A_{n+1} \text{ for } n > p,$$

or

$$(2)\ A_n^2 \leqq A_{n-1}A_{n+1} \text{ for } n > p.$$

PROOF. Condition (1) of Theorem I is satisfied, since

$$A_n^{-1} \sum_{k=1}^{n} a_{n-k+1} = 1.$$

We have

$$\sum_{n=1}^{\infty} \left| \sum_{p=k}^{\infty} (a_{np} - a_{n-1,p}) \right| = \sum_{n=k}^{\infty} | A_n^{-1} A_{n-k+1} - A_{n-1}^{-1} A_{n-k} |$$

where $a_0 = A_0 = 0$. If (1) is satisfied, we have

$$\begin{aligned}\sum_{n=k}^{m} | A_n^{-1} A_{n-k+1} - A_{n-1}^{-1} A_{n-k} | &\leqq \sum_{n=k}^{p+k-1} [| A_n^{-1} A_{n-k+1} | + | A_{n-1}^{-1} A_{n-k} |] \\ &\qquad + \sum_{n=p+k}^{m} (A_n^{-1} A_{n-k+1} - A_{n-1}^{-1} A_{n-k}) \\ &< 2p + A_m^{-1} A_{m-k+1} - A_{p+k-1}^{-1} A_p \\ &< 2(p + 1).\end{aligned}$$

It follows that condition (2) is satisfied. The second part of the theorem may be proved similarly.

COROLLARY 1. *A sufficient condition for the absolute regularity of the Nörlund mean A, where $a_n \geqq 0$, is the existence of a finite constant p such that*

$$a_n \geqq a_{n+1} \text{ for } n > p.$$

COROLLARY 2. *A sufficient condition for the absolute regularity of the Nörlund mean A, where $a_n \geqq 0$, is the existence of a finite constant p, such that*

$$a_n A_n \geqq a_{n+1} A_{n-1},$$

or

$$a_n A_n \leqq a_{n+1} A_{n-1}.$$

That the condition of Theorem II is not necessary is proved by the absolutely regular Nörlund mean A, for which

$$3 \cdot 2^n A_n = [2^n - (-1)^n] A_{n+1}.$$

It is easily verified by means of Theorem II that (C, δ), the Cesàro mean of order δ, is absolutely regular for $\delta \geqq 0$.[6]

If the matrices $\| a_{nk} \|$ and $\| b_{nk} \|$ are triangular, A includes B absolutely provided there exists an absolutely regular C such that $A = CB$. If B has an inverse, this condition is necessary and sufficient.

The following example shows that for non-triangular matrices, $A = CB$, where C is absolutely regular, does not imply that A includes B absolutely.

For $p \geqq 0$,

$$c_{nk} = \begin{cases} 2^{-k},\ k = 4\,p + 1;\ n = 1, 3, \cdots, 2p - 1;\ n \geqq 2p + 1; \\ 2^{-k},\ k = 4\,p + x;\ x = 2, 3, 4;\ n = 1, 3, \cdots, 2p + 1; \\ 2^{-k},\ k = 4\,p + 4;\ n \geqq 2p + 3; \\ -2^{-k}, \text{ all other } n; \end{cases}$$

and

$$b_{nk} = \begin{cases} 1 \quad\ ,\ n = 2,\ k = 1; \\ 4/3 \quad ,\ n \geqq 3,\ k = n - 1; \\ -1/3,\ n \geqq 1,\ k = n + 1; \\ 0 \quad\ ,\ \text{all other } k \text{ and } n. \end{cases}$$

The sequence $\{s_k\}$, $s_k = 2^k$, is transformed into an absolutely convergent series by B, but into a divergent series by CB, although C is absolutely regular.

3. We shall consider the Nörlund means A and B, defined by $\{a_n\}$ and $\{b_n\}$, sequences of complex numbers, such that $A_n \neq 0$, $B_n \neq 0$, for all n, and the sequence to sequence transformation defined by the matrix $\| a_{nk} \|$, where $a_{nk} = 0$ for $k > n$. We let $\{x_n\}$ and $\{y_n\}$ be sequences of complex numbers.

We define the transformation B' by $b_{nk} = B_n^{-1} b_k$ for $k \leqq n$, $b_{nk} = 0$ for $k > n$. We define the transformation T_l by $t_{nk} = a_{l+n-1,\, l+n-k}$ for $k \leqq n$, $t_{nk} = 0$ for $k > n$, where $l = 1, 2, \cdots$.

For any absolutely regular transformation, we shall call the greatest lower bound of constants which satisfy condition (2) of Theorem I, the *test constant* of the matrix of the transformation.

For the proof of Theorem III, we require the following lemma:

LEMMA. *If the series corresponding to* $\{x_n\}$ *and to* $\{B_n^{-1} \sum_{k=1}^{n} b_k y_k\}$ *converge absolutely, then the series corresponding to* $\{\sum_{k=1}^{n} a_{nk} x_k y_{n-k+1}\}$ *converges absolutely, provided*

(1) *there exists an absolutely regular* T_l' *such that* $T_l = T_l' B'$, $l = 1, 2, \cdots$;

(2) $M_l < M$, *where* M_l *is the test constant for* T_l', $l = 1, 2, \cdots$.

PROOF. Let $z_n = \sum_{k=1}^{n} c_{nk} x_k$ where $c_{nk} = a_{nk} y_{n-k+1}$. Since $\sum_{n=1}^{\infty} | x_n - x_{n-1} |$ converges, $\sum_{n=1}^{\infty} | z_n - z_{n-1} |$ converges if the transformation defined by c_{nk} is

[6] Kogbetliantz, loc. cit., p. 25.

absolutely regular. Condition (1) of Theorem I is satisfied. In order to satisfy condition (2) of Theorem I it is necessary that there exist a positive constant C such that

$$\sum_{n=1}^{\infty}\left|\sum_{p=k}^{\infty}(c_{np}-c_{n-1,p})\right| = \sum_{n=k}^{\infty}\left|\sum_{p=k}^{n}(a_{np}y_{n-p+1}-a_{n-1,p}y_{n-p})\right| \leqq C,$$

for all k.

Let $C = MY$, where $Y = \sum_{n=1}^{\infty} | B_n^{-1} \sum_{k=1}^{n} b_k y_k - B_{n-1}^{-1} \sum_{k=1}^{n-1} b_k y_k |$. For any k, say $k = k_1$,

$$\sum_{n=k}^{\infty}\left|\sum_{p=k}^{n}(a_{np}y_{n-p+1}-a_{n-1,p}y_{n-p})\right|$$

converges if T_{k_1} sums $\{y_n\}$ absolutely, that is, if T_{k_1} includes B' absolutely. By condition (1), T_{k_1} includes B' absolutely; by Theorem I,

$$\sum_{n=k}^{\infty}\left|\sum_{p=k}^{n}(a_{np}y_{n-p+1}-a_{n-1,p}y_{n-p})\right| \leqq M_{k_1}Y.$$

By condition (2), $M_kY < MY$ for all k. Therefore, it follows from Theorem I that $\sum_{n=1}^{\infty} | z_n - z_{n-1} |$ converges, which completes the proof.

If $C_n = \sum_{k=1}^{n} A_k b_{n-k+1} \neq 0$ for all n, the sequence $\{c_n\}$, where $c_n = \sum_{k=1}^{n} a_k b_{n-k+1}$, defines a Nörlund mean C. Let $\sum_{n=1}^{\infty} w_n$ be the Cauchy product of the series $\sum_{n=1}^{\infty} u_n$ and $\sum_{n=1}^{\infty} v_n$; let $W_k = \sum_{n=1}^{k} w_n$.

THEOREM III. *If $\sum_{n=1}^{\infty} u_n$ and $\sum_{n=1}^{\infty} v_n$ are absolutely summable A and B respectively, then $\sum_{n=1}^{\infty} w_n$ is absolutely summable C, provided C includes absolutely both A and B.*

PROOF. We have

$$C_n^{-1}\sum_{k=1}^{n} c_k W_{n-k+1} = \sum_{k=1}^{n}\left[C_n^{-1}A_k b_{n-k+1}A_k^{-1}\sum_{p=1}^{k}A_{k-p+1}u_p b_{n-k+1}^{-1}\sum_{p=1}^{n-k+1} b_{n-k-p+2}v_p\right].$$

Consider the triangular matrix $\| a_{nk} \|$, where $a_{nk} = C_n^{-1} A_k b_{n-k+1}$. For the transformation T_l, defined at the beginning of paragraph 3, we have

$$t_{nk} = \begin{cases} C_{l+n-1}^{-1} A_{l+n-k} b_k, & k \leqq n; \\ 0 & , k > n. \end{cases}$$

We have $T_l = T_l' B'$, where

$$t'_{nk} = \begin{cases} C_{l+n-1}^{-1} a_{l+n-k} B_k, & k < n; \\ C_{l+n-1}^{-1} A_l B_n & , k = n: \\ 0 & , k > n. \end{cases}$$

We define

$$A_x(p) = \frac{a_{x+n-1-p}}{C_{x+n-2}} - \frac{a_{x+n-p}}{C_{x+n-1}}.$$

we define $B_x(p)$ similarly in terms of b. We let $\sum_{p=1}^{x} f(p) = 0$, for $x < 1$ and $f(p)$ any function of p; we let $A_x = B_x = 0$, $x < 1$.

T_l satisfies condition (1) of the lemma, since T_l' is absolutely regular. For

$$\sum_{n=k}^{\infty} \left| \sum_{p=k}^{n} (t_{np}' - t_{n-1,p}') \right| = |R_1| + \sum_{n=k+1}^{\infty} |R_2|,$$

where

$$R_1 = \frac{A_l B_k}{C_{l+k-1}},$$

$$R_2 = B_n \frac{a_l}{C_{l+n-1}} - \sum_{p=k}^{n-1} B_p A_l(p) - A_{l-1}\left(\frac{B_{n-1}}{C_{l+n-2}} - \frac{B_n}{C_{l+n-1}}\right).$$

Since

$$A_l B_n = C_{l+n-1} - \sum_{p=1}^{l-1} A_p b_{l+n-p} - \sum_{p=1}^{n-1} B_p a_{l+n-p},$$

we have the following equalities:

$$R_1 = \sum_{n=1}^{k} \sum_{p=1}^{l-1} A_p B_l(p) + \sum_{n=1}^{l} \sum_{p=1}^{k-1} B_p A_k(p) + 1 - \operatorname{sgn}|l-1| - \operatorname{sgn}|k-1|;$$

$$\begin{aligned} R_2 &= \sum_{p=1}^{l-2} A_p B_l(p+1) + \sum_{p=1}^{l-1} a_p B_l(p) - \sum_{p=k}^{n-1} B_p A_l(p) \\ &\qquad + \sum_{p=1}^{n-2} B_p A_l(p+1) + \sum_{p=1}^{n-1} b_p A_l(p) \\ &= \sum_{p=1}^{l-1} A_p B_l(p) + \sum_{p=1}^{k-1} B_p A_l(p). \end{aligned}$$

We have

$$\sum_{n=k+1}^{\infty} \left| \sum_{p=1}^{k-1} B_p A_l(p) \right| = \sum_{n=l+1}^{\infty} \left| \sum_{p=1}^{k-1} B_p A_k(p) \right|.$$

Therefore

$$\begin{aligned} \sum_{n=k}^{\infty} \left| \sum_{p=k}^{n} (t_{np}' - t_{n-1,p}') \right| &\leq 1 + \sum_{n=1}^{\infty} \left| \sum_{p=1}^{l-1} A_p B_l(p) \right| + \sum_{n=1}^{\infty} \left| \sum_{p=1}^{k-1} B_p A_k(p) \right| \\ &= 1 + \sum_{n=1}^{\infty} |S_a| + \sum_{n=1}^{\infty} |S_b|. \end{aligned}$$

Since by hypothesis C includes A absolutely, $C = C'A$, where C' is absolutely regular, and there exists a constant M_a, such that

$$\sum_{n=1}^{\infty} \left| \sum_{p=l}^{\infty} (c_{np}' - c_{n-1,p}') \right| \leq M_a \text{ for all } l;$$

the same inequality holds for all k, since the left side is independent of k. But

$$\sum_{n=1}^{\infty}\left|\sum_{p=l}^{\infty}(c'_{np}-c'_{n-1,p})\right| = \begin{cases} 1 & , l=1; \\ \sum_{n=1}^{\infty}|S_a|, & l>1. \end{cases}$$

Similarly, $C = C''B$, and there exists a constant M_b such that

$$\sum_{n=1}^{\infty}\left|\sum_{p=k}^{\infty}(c''_{np}-c''_{n-1,p})\right| \leqq M_b \text{ for all } k \text{ and } l.$$

But

$$\sum_{n=1}^{\infty}\left|\sum_{p=k}^{\infty}(c''_{np}-c''_{n-1,p})\right| = \begin{cases} 1 & , k=1; \\ \sum_{n=1}^{\infty}|S_b|, & k>1. \end{cases}$$

Hence

$$\sum_{n=k}^{\infty}\left|\sum_{p=k}^{n}(t'_{np}-t'_{n-1,p})\right| \leqq 1 + M_a + M_b = M$$

for all k and l, and T'_l is absolutely regular for all l.

Condition (2) of the lemma is satisfied also, since M is independent of l.

The theorem follows immediately from the lemma.

It is obvious that Theorem III includes Cauchy's theorem for the multiplication of absolutely convergent series. It includes also as a special case the Kogbetliantz theorem for the Cesàro mean. For if $A = (C, \gamma)$ and $B = (C, \delta)$, where (C, x) represents the Cesàro mean of order x, $x \geqq 0$,[7] then $C = (C, \gamma + \delta)$. Since $(C, \gamma + \delta)$ includes (C, γ) and (C, δ) absolutely[8] the conditions of Theorem III are satisfied.

THE GEORGE WASHINGTON UNIVERSITY.

---

[7] Although Kogbetliantz states no limitation on γ and δ in his generalization of Cauchy's theorem, nor on γ in his generalization of Mertens' theorem (Kogbetliantz, loc. cit., p. 27), it is obvious that the theorems are true only if $\gamma \geqq 0$, $\delta \geqq 0$. For if we choose $\sum_{n=1}^{\infty} u_n$ absolutely summable $(C, \delta = 1)$, and $\sum_{n=1}^{\infty} v_n$ absolutely summable $(C, \gamma = -\epsilon)$, $\epsilon > 0$, we have $\gamma + \delta = 1 - \epsilon$; if we choose $v_1 = 1$, $v_n = 0$ for $n > 1$, $\sum_{n=1}^{\infty} w_n = \sum_{n=1}^{\infty} u_n$. But it is possible to construct a series, $\sum_{n=1}^{\infty} u_n$, which is absolutely summable $(C, 1)$, and not summable $(C, 1 - \epsilon)$, however small ϵ may be. (Kogbetliantz, loc. cit., p. 28.)

[8] Kogbetliantz, loc. cit., p. 25.

ANNALS OF MATHEMATICS
Vol. 38, No. 3, July, 1937

SPECIAL REGIONS OF REGULARITY OF FUNCTIONS OF SEVERAL COMPLEX VARIABLES[1]

Received December 4, 1935.

BY W. T. MARTIN[2]

1. **Introduction.** In the theory of functions of several complex variables various types of regions occur as convergence regions of series developments of analytic functions, notable ones being *Reinhardt regions* of absolute convergence of power-series, *circular regions* of uniform convergence of "diagonal" series (*the invariant convergence regions*) and *Cartan regions* of uniform convergence of certain other series (see sections 2–5). By associating with an analytic function $f(x, y)$ a class of functional transforms and by considering the growth of each of these transforms in certain types of *analytic surfaces* we obtain in part I the classic circular and Cartan regions related to the function $f(x, y)$. By a closer investigation of these transforms (along certain types of curves) we obtain in part II analytic continuations of the function $f(x, y)$ into new regions extending beyond the circular and Cartan regions. Analytic continuations of the function $f(x, y)$ are obtained also in part III by a consideration of the absolute summability by *Borel's integral means* of certain series developments of $f(x, y)$. The two methods of analytic continuation studied in parts II and III respectively are intimately related, as we show in part III. Also in part III we define and obtain properties of a class of regions, each region of which is an analogue of the *Mittag-Leffler star* of a function of a complex variable.

Most of the results mentioned in the previous paragraph are capable of furnishing a geometric interpretation of the growth of a class of entire functions, namely the class of entire functions consisting of the functional transforms used to obtain the results mentioned above. For the most part we consider two problems together, one as a problem relating to the investigation of functions regular at the origin, the other as a problem relating to the growth of entire functions. In the remainder of the introduction we shall indicate more explicitly these problems as well as the general nature of the results obtained in the paper.

[1] For a summary of the results of the present paper, see *Bulletin of the American Mathematical Society*, abstract Nos. 41-9-308, 42-1-50; *Proceedings of the National Academy of Sciences*, 22 (1936), pp. 33–35.

[2] National Research Fellow.

Let

$$f(x, y) = \sum_{m,n=0}^{\infty} a_{mn} x^m y^n \tag{1}$$

be any analytic function of the two complex variables x, y, regular at the origin. Let us consider the class of associated entire functions

$$F(x, y; p, q) = \sum_{m,n=0}^{\infty} \frac{a_{mn} x^m y^n}{(pm + qn)!}, \qquad (p, q \text{ positive integers}), \tag{2}$$

formed with the same constants a_{mn}. In part I we define a nonnegative valued function $H(x, y; p, q)$ which, for fixed (x_0, y_0), measures the increase of the function $F(x, y; p, q)$ in the analytic surface

$$x = x_0 z^p, \qquad y = y_0 z^q, \qquad (z \text{ a complex parameter}).$$

The set of points (x, y) for which $H(x, y; p, q) < 1$ is identical with the *complete Cartan (p, q)-region G_{pq} of uniform convergence* of the series[3]

$$f(x, y) = \sum_{\nu=0}^{\infty} \left\{ \sum_{pm+qn=\nu} a_{mn} x^m y^n \right\}. \tag{3}$$

A more precise indication of the growth of the function $F(x, y; p, q)$ is obtained in part II in terms of a real-valued function $h(x, y; p, q)$ which, for fixed (x_0, y_0), measures the increase of the function $F(x, y; p, q)$ along the curve

$$x = x_0 \rho^p, \qquad y = y_0 \rho^q, \qquad (\rho \text{ a positive parameter}).$$

By using the function $h(x, y; p, q)$ we obtain more information concerning the region of analyticity of the function $f(x, y)$. In terms of $h(x, y; p, q)$ we define a region D_{pq} which we call the *(p, q)-diagram* of $f(x, y)$, which contains the complete Cartan (p, q)-region G_{pq} and which furnishes an extended region of analyticity of $f(x, y)$.

In part III we investigate the absolute summability of the series (3) by Borel's integral means[4] and obtain the *Borel (p, q)-region B_{pq} of summability* of $f(x, y)$ which is such that $G_{pq} \subseteq B_{pq} \subseteq D_{pq}$. With B_{pq} we associate a region $\mathfrak{B}_{pq} \supseteq B_{pq}$, in which $f(x, y)$ is regular. In the case where $h(x, y; p, q)$ is everywhere non-negative we have $\mathfrak{B}_{pq} \equiv D_{pq}$.

In section 11 of part III we study, for every pair of positive numbers σ, μ, the range of analyticity of $f(x, y)$ along each of the curves

[3] For a discussion of Cartan regions see H. Cartan, "Les fonctions de deux variables complexes et le problème de la représentation analytique," *Journal de Mathématiques*, (9) 10 (1931), pp. 1–114; H. Behnke and P. Thullen, *Theorie der Funktionen mehrerer komplexer Veränderlichen*, Ergebnisse der Mathematik und ihrer Grenzgebiete, Berlin (1934), especially pp. 35, 43. See also section 5 where we state the definitions and properties essential for the remainder of the paper.

[4] For a discussion of summability by Borel's integral means see, for example, E. Borel, *Leçons sur les séries divergentes*, Paris, 2nd edition (1928), especially pp. 120–179; P. Dienes, *The Taylor series*, Oxford (1931), especially pp. 302–305.

$$x = \alpha\rho^{\sigma}, \qquad y = \beta\rho^{\mu}, \qquad (\rho \text{ a positive parameter}),$$

where α and β are any complex numbers such that $|\alpha| + |\beta| \neq 0$. We obtain a "star-shaped" region $S_{\sigma\mu}$ which we call the *(σ, μ)-star region* of $f(x, y)$ and which is analogous to the Mittag-Leffler star[5] of a function $f(z)$. By means of the study of the behavior of the functions

$$F_a(x, y; \sigma, \mu) = \sum_{m,n=0}^{\infty} \frac{a_{mn} x^m y^n}{\Gamma(1 + a\sigma m + a\mu n)}, \qquad (0 < a < \infty),$$

using the results of Riesz on Dirichlet series,[6] we define a real-valued function $c(x, y; \sigma, \mu)$ which is such that the (σ, μ)-star region $S_{\sigma\mu}$ is identical with the set of points (x, y) for which $c(x, y; \sigma, \mu) < 0$. The "vertices" of $S_{\sigma\mu}$ consist of those points (x, y) for which $c(x, y; \sigma, \mu) = 0$.

Pincherle, Borel, Pólya and others[7] have obtained information concerning regions of analyticity of a function $f(z) = \sum_0^\infty \alpha_\nu z^\nu$ by means of a study of the growth of an associated entire function $F(z) = \sum_{\nu=0}^\infty \alpha_\nu z^\nu/\nu!$. Corresponding to the fact that the *type* H of $F(z)$ indicates that the function $f(z)$ has the circle $|z| H < 1$ as the circle of convergence of its power-series development, we have the result that the *type-function* $H(x, y; p, q)$ of $F(x, y; p, q)$ is such that the function $f(x, y)$ has the complete Cartan (p, q)-region $H(x, y; p, q) < 1$ as the region of uniform convergence of its series development (3). The *indicator* $h(x, y; p, q)$ is an analogue of the *Phragmen-Lindelof indicator* $h(\theta)$; indeed the function $h(xe^{ip\theta}, ye^{iq\theta}; p, q)$ is a function of the Phragmen-Lindelof type.[8] The (p, q)-diagram D_{pq} is an analogue of the *indicator-diagram* of a function $F(z)$, which in its turn is related to the *Borel polygon of summability*.[9]

The methods of the entire paper are of such a nature that the extension to the theory of functions of more than two complex variables is obvious.

---

[5] For a discussion of Mittag-leffler stars see, for example, Dienes, *op. cit.*,[4] p. 308. See also B. Almer, "Sur quelques problèmes de la théorie des fonctions analytiques de deux variables complexes," *Arkiv för Matematik, Astronomi och Fysik* 17 (1922), 1-70, where stars of the sort considered here are introduced.

[6] M. Riesz, "Sur la représentation analytique des fonctions définies par des séries de Dirichlet," *Acta Mathematica*, 35 (1912), pp. 253-270. See also V. Bernstein, *Leçons sur les progrès de la théorie des séries de Dirichlet*, Paris (1933), especially pp. 184-192.

[7] S. Pincherle, "Della trasformazione di Laplace e di alcune sue applicazione," *Memorie della R. Accademia delle Scienze di Bologna*, (4) 8 (1887), pp. 125-143; E. Borel, *op. cit.*;[4] G. Polya, "Untersuchungen über Lücken und Singularitäten von Potenzreihen," *Mathematische Zeitschrift*, 29 (1929), pp. 549-640, especially pp. 571-610. See also other references in Polya's paper.

[8] See E. Phragmén and E. Lindelöf, "Sur une extension d'un principe classique de l' analyse," *Acta Mathematica*, 31 (1908), pp. 381-406.

[9] See Pólya, *loc. cit*,[7] especially p. 586, where the relation of the indicator-diagram to the Borel polygon is given. The (p, q)-diagram D_{pq}, as we define it, is more nearly the analogue of the complement of the conjugate diagram defined by Polya. Just as the complement of the inverted conjugate diagram is related to the Borel polygon so the (p, q)-diagram D_{pq} is related to the Borel (p, q)-region of summability B_{pq}.

I am indebted to Professor Bochner for the suggestions he made during the preparation of this paper.

I. Type-Functions and Cartan Regions

2. Preliminary considerations. In sections 2–5 we will be concerned with a function-element

$$f(x, y) = \sum_{m,n=0}^{\infty} a_{mn} x^m y^n \tag{1}$$

of the two complex variables x, y, whose expansion point is taken without loss of generality as the point $x = 0, y = 0$. We shall discuss convergence regions and regions of analyticity of such a function-element; the cases where the function-element always converges or always (except for (0, 0)) diverges being without interest in this connection, we exclude these two possibilities. More complete discussions of these matters are given in various places.[10] We merely state here the definitions and properties which we shall later use.

3. Regions of absolute convergence. Reinhardt regions.

Definition 1. A pair of positive numbers (r_0, r_0') is called a *pair of associated convergence-radii* of the power-series (1) if this series converges absolutely for all (x, y) for which $|x| < r_0$, $|y| < r_0'$, but, on the contrary, diverges for all (x, y) for which $|x| > r_0$, $|y| > r_0'$.

1°. *If (r_0, r_0') is a pair of associated convergence-radii of the power-series* (1), *then we always have*

$$\limsup_{m+n=\infty} | a_{mn} r_0^m r_0'^n |^{\frac{1}{m+n}} = 1.$$

2°. *If (r_0, r_0') is a pair of associated convergence-radii of* (1), *then there exists at least one point (x_0, y_0) with $|x_0| = r_0$, $|y_0| = r_0'$, in which the function $f(x, y)$ represented by* (1) *is singular.*

Definition 2. By a *complete Reinhardt region* we understand a region which with a point (x_0, y_0) contains also all the points (kx_0, ly_0) with $|k| \leqq 1$, $|l| \leqq 1$.

3°. *The region of absolute and uniform convergence of the power-series* (1) *is a complete Reinhardt region R. If we set $|x| = r$, $|y| = r'$, then the pairs of associated convergence-radii of* (1) *constitute a three-dimensional manifold in the x-y-space—the boundary of the convergence region R.*

4. Invariant convergence regions. Circular regions.

Definition 3. By a *complete circular region* we mean a region which with a point (x_0, y_0) contains also all the points (kx_0, ky_0) with $|k| \leqq 1$.

---

[10] See, for example, Behnke and Thullen, *op. cit.*,[3] where also references to the original memoirs relating to these subjects are given. In general we shall follow the terminology adopted by Behnke and Thullen.

DEFINITION 4. By an *analytic plane* we mean a two-dimensional linear manifold which is represented by a complex equation of the form $ax + by + c = 0$.

4°. *Every analytic plane through the origin cuts a complete circular region in the interior of a circle.*

The convergence region R of the power-series development (1) of the function $f(x, y)$ depends upon the arbitrary choice of the coordinate planes. We form the set $\mathfrak{M}$ of all points $P(x, y)$ with the property that after a suitable linear transformation of coordinates

$$x' = ax + by, \qquad y' = cx + dy, \qquad ad - bc \neq 0,$$

the power-series $f(x(x', y'), y(x', y'))$, according to powers of x', y', converges in a neighborhood of P (considered in the x'-y'-system). The largest (open) region contained in $\mathfrak{M}$ is a *schlicht* region G which is called the *invariant convergence region* of the function $f(x, y)$.

5°. *The invariant convergence region G of a function $f(x, y)$ with the power-series development* (1) *is identical with the region of uniform convergence of the "diagonal" series*

$$f(x, y) = \sum_{\nu=0}^{\infty}\left\{\sum_{m+n=\nu} a_{mn}x^m y^n\right\}.$$

The region G is a complete circular region and is that particular one which possesses the property that every analytic plane through the origin cuts G in the interior of a circle whose circumference passes through the nearest singularity of $f(x, y)$ in the analytic plane considered.

5. Cartan regions.

DEFINITION 5. Let p and q be any two positive integers. By a *complete Cartan (p, q)-region* we understand a region which with a point (x_0, y_0) contains also all the points $(k^p x_0, k^q y_0)$ with $|k| \leqq 1$.

6°. *If a function $f(x, y)$ with the power-series development* (1) *is regular in a complete Cartan (p, q)-region G_{pq}, then the function $f(x, y)$ possesses a uniformly convergent expansion of the form* (3), *valid in the interior of G_{pq}.*

6. The type-function $H(x, y; p, q)$.

Let

$$f(x, y) = \sum_{m,n=0}^{\infty} a_{mn}x^m y^n \tag{1}$$

be any function regular at the origin. We exclude once and for all the case where $f(x, y)$ is identically zero. Let p and q be any two positive integers and let us define the *Borel (p, q)-transform* $F(x, y; p, q)$ of $f(x, y)$ in the following manner:

$$F(x, y; p, q) = \sum_{m,n=0}^{\infty} \frac{a_{mn}x^m y^n}{(pm + qn)!}, \qquad (p, q \text{ positive integers}). \tag{2}$$

In this section we shall obtain the region of uniform convergence of the series

$$f(x, y) = \sum_{\nu=0}^{\infty} \left\{ \sum_{pm+qn=\nu} a_{mn} x^m y^n \right\} \tag{3}$$

in terms of the behavior of the function $F(x, y; p, q)$.[11]

Using the results of section 3, we easily see that

7°. *The function $F(x, y; p, q)$ is an entire function.*

Let

$$M(\rho; x, y; p, q) = \max_{|z| \leq \rho} F(xz^p, yz^q, p, q) \tag{4}$$

and let us define the real-valued functions

$$H^*(x, y; p, q) = \limsup_{\rho=\infty} \frac{\log M(\rho; x, y; p, q)}{\rho}, \tag{5}$$

$$H(x, y; p, q) = \limsup_{x'=x,\, y'=y} H^*(x', y'; p, q). \tag{6}$$

We next state and prove the first theorem.

---

[11] L. Baumgartner, "Beiträge zur Theorie der ganzen Funktionen von zwei komplexen Veränderlichen," *Monatshefte für Mathematik und Physik*, 25 (1914), pp. 3–70, has investigated properties of certain classes of entire functions. One class which he has considered is the class of entire functions defined by series of the form

$$\sum_{m,n=0}^{\infty} \frac{a_{mn} x^m y^n}{m!\,n!}, \tag{α}$$

where the constants a_{mn} are such that the series (1) converges in a neighborhood of the origin. The results obtained by Baumgartner for functions of this form may be interpreted as giving information about the pairs of associated convergence-radii of the power-series (1). For the purpose of obtaining analytic continuations of $f(x, y)$ it seems more suitable to associate with $f(x, y)$ functions defined by series of the form (2) rather than functions defined by series of the form

$$\sum_{m,n=0}^{\infty} \frac{a_{mn} x^m y^n}{(pm)!\,(qn)!}. \tag{β}$$

Let us also discuss what happens if we permit p and q to be arbitrary positive numbers. If we group together, as in (3), the terms for which $pm + qn$ has the same value, then we have two distinct situations. (a) If p and q have a rational ratio, then the series corresponding to (3) is identical with the series

$$\sum_{\nu=0}^{\infty} \left\{ \sum_{p'm+q'n=\nu} a_{mn} x^m y^n \right\}, \tag{γ}$$

where p', q' are positive integers such that $(p/q) = (p'/q')$. (b) If p and q have an irrational ratio, then the series corresponding to (3) is a simple series in which the bracketed expression always contains a single term. In this case it is easily seen that the region of uniform convergence of such a series is identical with the Reinhardt region R of the power-series (1). Consequently in parts I and II we consider only the cases when p and q are positive integers.

THEOREM 1. *Let $f(x, y)$ be any function regular at the origin and let* (1) *be its power-series development. Let $F(x, y; p, q)$ and $H(x, y; p, q)$ be defined as in* (2) *and* (6) *respectively. Then the region G_{pq} of uniform convergence of the series* (3) *is identical with the set of points (x, y) for which $H(x, y; p, q) < 1$. The region G_{pq} is a complete Cartan (p, q)-region and is that particular one which possesses the property that every analytic surface of the form*

$$(7) \qquad x = \alpha z^p, \qquad y = \beta z^q, \qquad (|\alpha| + |\beta| \neq 0, \; z \text{ a complex parameter}),$$

cuts it in a two-dimensional region

$$(8) \qquad x = \alpha z^p, \qquad y = \beta z^q, \qquad |z| < 1/H(\alpha, \beta; p, q),$$

whose boundary is a closed curve

$$(9) \qquad x = \alpha z^p, \qquad y = \beta z^q, \qquad |z| = 1/H(\alpha, \beta; p, q),$$

which passes through the nearest singularity of $f(x, y)$ in the analytic surface (7).

From the results of Cartan and Hartogs[12] it is clear that the region G_{pq} of uniform convergence of (3) is identical with the set of points (x, y) for which $H_+(x, y; p, q) < 1$, where

$$(10) \qquad H_+^*(x, y; p, q) = \limsup_{\nu=\infty} \Big| \sum_{pm+qn=\nu} a_{mn} x^m y^n \Big|^{1/\nu},$$

$$(11) \qquad H_+(x, y; p, q) = \limsup_{x'=x,\, y'=y} H_+^*(x', y'; p, q).$$

We shall next prove that

$$(12) \qquad H_+(x, y; p, q) = H(x, y; p, q).$$

In order to prove (12) we use the following result from the theory of functions of one complex variable:[13]

LEMMA 1. *Let*

$$F(z) = \sum_{\nu=0}^{\infty} \frac{\alpha_\nu z^\nu}{\nu!}$$

be any function for which

$$\limsup_{\nu=\infty} |\alpha_\nu|^{1/\nu} = \eta_+ < \infty .$$

Let $M(\rho)$ denote the maximum of $|F(z)|$ for $|z| \leqq \rho$ and let

$$\limsup_{\rho=\infty} \frac{\log M(\rho)}{\rho} = \eta .$$

---

[12] See Cartan, *loc. cit.*;[3] F. Hartogs, "Zur Theorie der analytischen Funktionen mehrerer unabhängiger Veränderlichen, etc.," *Mathematische Annalen*, 62 (1906), pp. 1–88, especially pp. 1–15.

[13] For the proof of Lemma 1 see, for example, Pólya, *loc. cit.*,[7] pp. 578–580.

Then

$$\eta_+ = \begin{cases} 0, \text{ if } \eta = -\infty, \\ \eta, \text{ if } \eta \neq -\infty. \end{cases}$$

The case $\eta = -\infty$ occurs if and only if $F(z)$ is identically zero.

Applying Lemma 1 to the function

$$F(xz^p, yz^q; p, q) = \sum_{\nu=0}^{\infty} \frac{z^\nu}{\nu!} \left\{ \sum_{pm+qn=\nu} a_{mn} x^m y^n \right\},$$

considered as a function of z for fixed x, y, and recalling (10) and (5), we see that

$$H_+^*(x, y; p, q) = \begin{cases} 0, \text{ if } H^*(x, y; p, q) = -\infty, \\ H^*(x, y; p, q), \text{ if } H^*(x, y; p, q) \neq -\infty.\text{[14]} \end{cases} \tag{13}$$

From (6), (11) and (13) we see that (12) is true except possibly when $H(x, y; p, q) = -\infty$. But the situation $H(x, y; p, q) = -\infty$ can never occur since it is easily seen to imply that $F(x, y; p, q)$ is identically zero and this is impossible since $f(x, y)$ is not identically zero. Hence (12) always holds. This concludes the proof of the first assertion made in Theorem 1.

That G_{pq} is a complete Cartan (p, q)-region follows readily from Cartan's theory or we may see it by noting that the function $H(x, y; p, q)$ possesses the following property:

$$H(k^p x, k^q y; p, q) = |k| \, H(x, y; p, q), \qquad |k| < \infty. \tag{14}$$

For the proof of the remaining portion of the theorem we observe that by (14) an analytic surface of the form (7) cuts G_{pq} in the set of points (x, y) in (8) and that this point set is a two-dimensional region whose boundary is the closed curve defined by (9). We next show that the function $f(x, y)$ has a singularity on the curve (9). The function

$$f(\alpha' z^p, \beta' z^q) = \sum_{\nu=0}^{\infty} z^\nu \left\{ \sum_{pm+qn=\nu} a_{mn} \alpha'^m \beta'^m \right\}, \tag{15}$$

considered as a function of z, has a singularity at some point $z = z_{\alpha'\beta'}$, where

$$|z_{\alpha'\beta'}| = \frac{1}{H_+^*(\alpha', \beta'; p, q)}. \tag{16}$$

Let (α_s, β_s) be a sequence of points converging to (α, β) and such that

$$\lim_{s=\infty} z_{\alpha_s \beta_s} = z_0, \qquad \lim_{s=\infty} H_+^*(\alpha_s, \beta_s; p, q) = H(\alpha, \beta; p, q). \tag{17}$$

[14] It is easily verified that $H_+^*(x, y; p, q)$ is always finite. For, since $f(x, y)$ is regular at the origin, $H_+^*(x, y; p, q) \leq 1$ in a neighborhood of the origin. Furthermore we clearly have $H_+^*(k^p x, k^q y; p, q) = |k| \, H_+^*(x, y; p, q)$, $|k| < \infty$. These two facts show that $H_+^*(x, y; p, q)$ is always finite.

From (11) and (12) we see that such a sequence exists. The function $f(x, y)$ cannot be regular at any of the points $(x = \alpha_s z^p_{\alpha_s\beta_s},\ y = \beta_s z^q_{\alpha_s\beta_s})$ and consequently it cannot be regular at the limit point $(x = \alpha z_0^p,\ y = \beta z_0^q)$, which certainly is a point on the curve (9).[15]

This concludes the proof of Theorem 1.

Using (14) and the results contained in Theorem 1 concerning the presence of singularities on the boundary of G_{pq}, we derive the following property of $H(x, y; p, q)$:

8°. *The type-function* $H(x, y; p, q)$ *is upper semi-continuous; indeed,*

$$\limsup_{x=x_0,\, y=y_0} H(x, y; p, q) = H(x_0, y_0; p, q).^{16} \tag{18}$$

We omit the details of the proof.

We next prove the following theorem relating to the case when $H(x, y; p, q)$ is continuous.

THEOREM 2. *Let* $F(x, y; p, q)$ *be any entire function of the form* (2) *which is such that the associated series* (1) *converges in a neighborhood of the origin and which is such that the type-function* $H(x, y; p, q)$ *defined in* (6) *is continuous. Then for every positive number* ϵ *there exists a constant* A_ϵ, *dependent on* ϵ *in general but independent of* x, y, *such that we have*

$$| F(x, y; p, q) | < A_\epsilon e^{H(x,y;p,q)+\epsilon|x|^{1/p}+\epsilon|y|^{1/q}} \tag{19}$$

for every finite point (x, y).

In order to prove Theorem 2 let $F(x, y; p, q)$ be any function for which the hypotheses of the theorem are satisfied. Let ϵ be any positive number and let us consider the closed region S_ϵ which consists of the set of points (x, y) for which

$$H(x, y; p, q) + \epsilon \mid x \mid^{1/p} + \epsilon \mid y \mid^{1/q} \leqq 1.^{17} \tag{20}$$

Since S_ϵ is a closed region contained entirely within G_{pq} there exists a positive quantity A_ϵ such that

$$| f(x, y) | < A_\epsilon \tag{21}$$

[15] Let us remark that an alternative proof of the existence of a singularity on each of the curves (9) can be given by a method analogous to that used by Behnke in proving the corresponding result for convergence regions of diagonal series. See H. Behnke, "Natürliche Grenzen," *Hamburgischen Universität Abhandlungen*, 5 (1926–1927), pp. 290–312, especially pp. 300–301. Behnke's results also show that for the case $p = q = 1$ the region G_{11}, which is a complete circular region, is the invariant convergence region of $f(x, y)$.

[16] For the case $p = q = 1$ the result contained in 8° is equivalent to a classical result due to Hartogs, *loc. cit.*[12]

[17] The set of points represented by (20) is closed since $H(x, y; p, q)$ is assumed to be continuous.

for every point (x, y) in S_ϵ. It is easily verified that $F(x, y; p, q)$ is representable in the form

$$F(x, y; p, q) = \frac{1}{2\pi i} \int_{C_{\epsilon; x, y}} e^z f\left(\frac{x}{z^p}, \frac{y}{z^q}\right) \frac{dz}{z}, \tag{22}$$

where $C_{\epsilon; x, y}$ is the circle $| z | = H(x, y; p, q) + \epsilon | x |^{1/p} + \epsilon | y |^{1/q}$ and (x, y) is any finite point other than the origin. Obviously for every point (x, y) $\neq (0, 0)$ and for every point z on $C_{\epsilon; x, y}$, we have

$$\left| f\left(\frac{x}{z^p}, \frac{y}{z^q}\right) \right| < A_\epsilon. \tag{23}$$

Applying (23) to (22), we see that (19) holds except possibly at (0, 0). But

$$| F(0, 0; p, q) | = | f(0, 0) | < A_\epsilon$$

and hence (19) holds also for (0, 0).

II. Indicator-Functions and Diagrams

7. **Functions of exponential type of one complex variable.** We shall give here a brief resume of the parts of Pólya's investigations which we shall use later in this part.[18]

DEFINITION 6. A function

$$F(z) = \sum_{n=0}^{\infty} \frac{a_n z^n}{n!} \tag{24}$$

is of *exponential type* if and only if the series

$$\frac{1}{z} f\left(\frac{1}{z}\right) = g(z) = \sum_{n=0}^{\infty} \frac{a_n}{z^{n+1}} \tag{25}$$

converges in a neighborhood of $z = \infty$.

In this section we shall consider a function $F(z) \not\equiv 0$ of exponential type and its associated function $g(z)$.

9°. *If $M(\rho)$ is the maximum of $| F(z) |$ for $| z | \leqq \rho$, then*

$$H = \limsup_{\rho=\infty} \frac{\log M(\rho)}{\rho} = \limsup_{n=\infty} | a_n |^{1/n}. \tag{26}$$

DEFINITION 7. By the *indicator* $h(\theta)$ of $F(z)$ we mean the function

$$h(\theta) = \limsup_{\rho=\infty} \frac{\log | F(\rho e^{i\theta}) |}{\rho}. \tag{27}$$

By the *conjugate diagram* Δ of the function $F(z)$ we mean the set of points z defined as follows. A point $z = \rho e^{i\phi}$ belongs to Δ if and only if the inequality

[18] In this section we merely state certain definitions and results; details will be found in the article by Pólya, *loc. cit.*[7]

$$(28) \qquad \rho \cos(\theta + \phi) \leqq h(\theta)$$

is satisfied for every real value of θ.

10°. *A real-valued function* $h(\theta)$ *defined for all real values of the variable* θ *is the indicator of a function* $F(z)$ *of exponential type if and only if* (i) *it is periodic of period* 2π *and* (ii) *for all real values of* θ_1, θ_2, θ_3 *which fulfill the conditions*

$$(29) \qquad \theta_1 < \theta_2 < \theta_3, \qquad \theta_2 - \theta_1 < \pi, \qquad \theta_3 - \theta_2 < \pi,$$

it satisfies the inequality

$$(30) \qquad h(\theta_1)\sin(\theta_3 - \theta_2) + h(\theta_2)\sin(\theta_1 - \theta_3) + h(\theta_3)\sin(\theta_2 - \theta_1) \geqq 0.$$

11°. $-H \leqq h(\theta) \leqq H, \qquad \max_{0 \leqq \theta \leqq 2\pi} h(\theta) = H.$

12°. *The conjugate diagram* Δ *is a closed convex region contained in the circle* $|z| \leqq H$. *The function* $g(z)$ *is regular for every point* z *exterior to* Δ *and has every extreme point of* Δ *as a singular point.*[19] *If* Δ' *is any closed convex region which possesses the property that* $g(z)$ *is regular for every point* z *exterior to* Δ' *then* $\Delta' \supseteqq \Delta$.

13°. *For every closed convex region* Δ_1 *(in the bounded portion of the z-plane) there exists a function* $F_1(z)$ *of exponential type which has* Δ_1 *as its conjugate diagram.*[20]

8. **Diagrams.** Let $f(x, y)$ be any function regular at the origin and let $F(x, y; p, q)$ be its Borel (p, q)-transform defined as in (2). Let us define the real-valued functions

$$(31) \qquad h^*(x, y; p, q) = \limsup_{\rho=\infty} \frac{\log |F(x\rho^p, y\rho^q; p, q)|}{\rho},$$

$$(32) \qquad h(x, y; p, q) = \limsup_{x'=x, y'=y} h^*(x', y'; p, q).$$

Using the (p, q)-*indicator* $h(x, y; p, q)$, we define the (p, q)-*diagram* D_{pq} and we investigate the behavior of $f(x, y)$ in D_{pq}.

DEFINITION 8. By the (p, q)-*diagram* D_{pq} of $F(x, y; p, q)$ we mean the set of points (x, y) defined as follows.

(i) *A finite point* (x_0, y_0) *belongs to* D_{pq} *if and only if there exists a real value* θ_0 *such that*

$$(33) \qquad h(x_0 e^{ip\theta_0}, y_0 e^{iq\theta_0}; p, q) < \cos\theta_0 .$$

[19] By an *extreme point* of a convex region we understand a boundary point which is not an inner point of any straight line segment contained on the boundary.

[20] We shall refer to the indicator $h_1(\theta)$ of such a function $F_1(z)$ as the *indicator-function* associated with the convex region Δ_1. It is uniquely determined by the convex region Δ_1; indeed, $h_1(-\theta)$ is the classical *supporting function* of Δ_1. See, for example, Pólya, *loc. cit.*,[7] pp. 573-578. Lines of the form $\rho\cos(\theta + \varphi) = h_1(\theta)$ are *supporting-lines* of Δ_1. Through every boundary point of Δ_1 there passes at least one supporting-line of Δ_1. Every supporting-line of Δ_1 contains at least one extreme point.

(ii) *An infinite point with homogeneous coordinates*[21] $(x_1, y_1, 0)$ *belongs to* D_{pq} *if and only if there exists a real value* θ_1 *such that*

$$h(x_1 e^{ip\theta_1}, y_1 e^{iq\theta_1}; p, q) < 0. \tag{34}$$

Before proceeding with the investigation of the function $f(x, y)$ in D_{pq} we first derive the following useful properties of the (p, q)-indicator:

14°. $h(\rho^p x, \rho^q y; p, q) = \rho h(x, y; p, q),\ 0 \leqq \rho < \infty$;

15°. $-H(x, y; p, q) \leqq h(xe^{ip\theta}, ye^{iq\theta}; p, q) \leqq H(x, y; p, q),\ -\infty < \theta < \infty$;

16°. *For every pair of values* (α, β) *for which* $|\alpha| + |\beta| \neq 0$, *the function* $h(\alpha e^{ip\theta}; \beta e^{iq\theta}; p, q)$, *considered as a function of the real variable* θ, *is the indicator of a function* $F(z)$ *of exponential type. The set of points* $z = \rho e^{i\phi}$ *for which*

$$\rho \cos(\theta + \phi) \leqq h(\alpha e^{ip\theta}, \beta e^{iq\theta}; p, q)$$

for every real value of θ, *constitutes a closed convex region in the z-plane.*

17°. $h^*(x, y; p, q) \leqq h(x, y; p, q)$.

For the proof of 14° we observe that a similar relation with h replaced by h^* holds. From this fact and (32) we see that 14° holds.

Recalling Definition 6 and using the convergence of the power-series (1), we see that the function $F(xz^p, yz^q; p, q)$, considered as a function of z for fixed (x, y), is of exponential type with indicator $h^*(xe^{ip\theta}, ye^{iq\theta}; p, q)$. Hence, in view of 11°, we have

$$-H^*(x, y; p, q) \leqq h^*(xe^{ip\theta}, ye^{iq\theta}; p, q) \leqq H^*(x, y; p, q).$$

From (6) and (32) we see that 15° follows.

As we just observed, the function $h^*(\alpha e^{ip\theta}, \beta e^{iq\theta}; p, q)$, as a function of θ, is the indicator of a function of exponential type in z (namely, the function $F(\alpha z^p, \beta z^q; p, q)$). Hence each of the functions $h^*(\alpha e^{ip\theta}, \beta e^{iq\theta}; p, q)$, $|\alpha| + |\beta| \neq 0$, satisfies the conditions in 10°. From the defining relation (32) we easily see that this implies that each of the functions $h(\alpha e^{ip\theta}, \beta e^{iq\theta}; p, q)$, $|\alpha| + |\beta| \neq 0$, satisfies the conditions in 10°. This proves that each of the functions $h(\alpha e^{ip\theta}, \beta e^{iq\theta}; p, q)$ is the indicator of a function of exponential type in z. By Definition 7 and the first statement in 12° we see that the set of points z described in 16° is a closed convex region. This concludes the proof of 16°.

Noting that

$$\begin{aligned} h(x, y; p, q) &= \limsup_{x'=x,\, y'=y} h^*(x', y'; p, q) \geqq \lim_{\nu=\infty} h^*\left(x\left(1 + \frac{1}{\nu}\right)^p, y\left(1 + \frac{1}{\nu}\right)^q; p, q\right) \\ &= h^*(x, y; p, q), \end{aligned}$$

we see the validity of 17°.

---

[21] We follow the convention adopted by Behnke and Thullen concerning infinite points (*op. cit.*,[3] pp. 3-6). For example, we think of the x, y-space extended by an analytic plane at infinity.

We next prove the following lemma:

LEMMA 2. *The function $f(x, y)$ is regular at every point (x_0, y_0) which is expressible in the form*

$$x_0 = \alpha/z^p, \qquad y_0 = \beta/z^q,$$

where α, β are complex numbers and $z = \rho e^{i\phi}$ is any quantity different from zero and such that there exists a real value θ for which

$$h(\alpha e^{ip\theta}, \beta e^{iq\theta}; p, q) < \rho \cos(\theta + \varphi). \tag{35}$$

Let θ_0 be any real number and α_0, β_0 any complex numbers. From (6) and (32) we see that for every positive number ϵ there exists a neighborhood I_ϵ of (α_0, β_0) of such a sort that we have both

$$h^*(\alpha e^{ip\theta_0}, \beta e^{iq\theta_0}; p, q) < h(\alpha_0 e^{ip\theta_0}, \beta_0 e^{iq\theta_0}; p, q) + \epsilon \tag{36}$$

and

$$H^*(\alpha, \beta; p, q) < H(\alpha_0, \beta_0; p, q) + \epsilon, \tag{37}$$

for every point (α, β) in I_ϵ. Using (37) and recalling the results of Theorem 1, we obtain the result that the function $f(x, y)$ is regular at every point (x, y) which is expressible in the form

$$x = \alpha/z^p, \qquad y = \beta/z^q,$$

where (α, β) is in I_ϵ and where z is any complex number such that

$$| z | > H(\alpha_0, \beta_0; p, q) + \epsilon.$$

By classical theory[22] this implies that the function

$$g(\alpha, \beta, z) = \frac{1}{z} f\left(\frac{\alpha}{z^p}, \frac{\beta}{z^q}\right), \tag{38}$$

as a function of α, β, z, is regular for (α, β) in I_ϵ and z such that $| z | > H(\alpha_0, \beta_0; p, q) + \epsilon$. Our immediate goal is to show that the function $g(\alpha, \beta, z)$ is regular at every point $P_0(\alpha_0, \beta_0, z_0)$, where $z_0 = \rho_0 e^{i\varphi_0}$ is any value such that

$$h(\alpha_0 e^{ip\theta_0}, \beta_0 e^{iq\theta_0}; p, q) < \rho_0 \cos(\theta_0 + \varphi_0). \tag{39}$$

If z_0 is such that $| z_0 | > H(\alpha_0, \beta_0; p, q)$, then the function $g(\alpha, \beta, z)$ is already regular at P_0. If $| z_0 | \leqq H(\alpha_0, \beta_0; p, q)$ we proceed. Let ϵ_0 be a positive number such that

$$h(\alpha_0 e^{ip\theta_0}, \beta_0 e^{iq\theta_0}; p, q) + \epsilon_0 < \rho_0 \cos(\theta_0 + \varphi_0).$$

Let L be a half-ray in the z-plane which proceeds from the point z_0 perpendicular to the line

$$h(\alpha_0 e^{ip\theta_0}, \beta_0 e^{iq\theta_0}; p, q) + \epsilon_0 = \rho \cos(\theta_0 + \varphi) \tag{40}$$

[22] See, for example, W. F. Osgood, *Lehrbuch der Funktionentheorie*, vol. 2, Berlin (1924), p. 7.

and contained in the half-plane

$$h(\alpha_0 e^{ip\theta_0}, \beta_0 e^{iq\theta_0}; p, q) + \epsilon_0 < \rho \cos(\theta_0 + \varphi).$$

Let z_1 be any point on L exterior to the circle $|z| = H(\alpha_0, \beta_0; p, q) + \epsilon_0$ and and let δ be a positive number such that every point z' for which $|z' - z_1| < \delta$ is also exterior to the circle $|z| = H(\alpha_0, \beta_0; p, q) + \epsilon_0$. Then $g(\alpha, \beta, z)$ is regular at every point (α', β', z') for which (α', β') is in I_{ϵ_0} and z' is in the circle $|z - z_1| < \delta$.[23] Noting that for fixed (α, β) the function $F(\alpha z^p, \beta z^q; p, q)$ is a function of exponential type in z with indicator $h^*(\alpha e^{ip\theta}, \beta e^{iq\theta}; p, q)$ and associated function $g(\alpha, \beta, z)$, using (36) for $\epsilon = \epsilon_0$ and using property 12°, we see that $g(\alpha, \beta, z)$, considered as a function of z, is surely regular at every point z in the circle C which has as its center the point z_1 and which is tangent to the line (40), this being true for every pair of values (α, β) in I_{ϵ_0}. Let us denote by r the radius of the circle C. Then we have proved the following two facts: (i) the function $g(\alpha, \beta, z)$ is an analytic function of the three complex variables α, β, z, regular for (α, β) in I_{ϵ_0} and z in the circle $|z - z_1| < \delta$; (ii) the function $g(\alpha, \beta, z)$ is an analytic function of the complex variable z regular for z in the circle $|z - z_1| < r$ for every pair of values (α, β) in I_{ϵ_0}. From these two facts, using a classical result due to Hartogs,[24] we obtain the result that $g(\alpha, \beta, z)$ is an analytic function of the three variables α, β, z, regular at every point (α, β, z) for which (α, β) is in I_{ϵ_0} and $|z - z_1| < r$. Hence, in particular, $g(\alpha, \beta, z)$ is regular at the point $P_0(\alpha_0, \beta_0, z_0)$.

If z_0 is different from zero, the regularity of $g(\alpha, \beta, z) = \frac{1}{z} f\left(\frac{\alpha}{z^p}, \frac{\beta}{z^q}\right)$ at the point (α_0, β_0, z_0) implies the regularity of $f(\alpha/z_0^p, \beta/z_0^q)$, as a function of α, β, at the point (α_0, β_0), which is equivalent to the statement that $f(x, y)$ is regular at the point $(x = \alpha/z_0^p, y = \beta/z_0^q)$. Recalling that z_0 is any value of z for which (39) holds, we see that the proof of Lemma 2 is complete.

The result which we next obtain is used in the proof of Theorem 3.

LEMMA 3. *Let θ_0 be any real number and α_0, β_0, any complex numbers such that $|\alpha_0| + |\beta_0| \neq 0$. Then the function $g(\alpha, \beta, z)$ in* (38), *considered as a function of α, β, z, is regular in the half-plane*

$$E \qquad \alpha = \alpha_0, \beta = \beta_0, \rho\cos(\theta_0 + \varphi) > h(\alpha_0 e^{ip\theta_0}, \beta_0 e^{iq\theta_0}; p, q)$$

and it has a singularity on the line

$$L \qquad \alpha = \alpha_0, \beta = \beta_0, \rho\cos(\theta_0 + \varphi) = h(\alpha_0 e^{ip\theta_0}, \beta_0 e^{iq\theta_0}; p, q).$$

---

[23] See the remarks in the sentence containing equation (38).

[24] F. Hartogs, *loc. cit.*,[12] especially pp. 18–19. A statement of Hartogs' result in the form in which we shall use it in the present paper is given in a paper by S. Bochner and W. T. Martin, "Singularities of composite functions in several variables," *Annals of Mathematics*, (2) 38 (1937), 293–303. (Theorem B, p. 294.)

In the process of proving Lemma 2 we proved that $g(\alpha, \beta, z)$ is regular at every point (α_0, β_0, z_0) where z_0 is any value satisfying (39). This implies that $g(\alpha, \beta, z)$ is regular in the half-plane E.

As a function of z alone, the function $g(\alpha', \beta', z)$ has a singularity on the *supporting-line*

$$\rho \cos(\theta_0 + \varphi) = h^*(\alpha' e^{ip\theta_0}, \beta' e^{iq\theta_0}; p, q)$$

(see 12° and footnote 20). Hence $g(\alpha, \beta, z)$, as a function of α, β, z, has a singular point on each of the lines

$$\alpha = \alpha', \qquad \beta = \beta', \qquad \rho \cos(\theta_0 + \varphi) = h^*(\alpha' e^{ip\theta_0}, \beta' e^{iq\theta_0}; p, q),$$

and hence by (33), it has a singularity on L.

We next prove the following theorem:

THEOREM 3. *Let $f(x, y)$ be any function regular at the origin and let* (1) *be its power-series development. Let $F(x, y; p, q)$, $h(x, y; p, q)$ and D_{pq} be defined as in* (2), (32) *and Definition* 8, *respectively. Then the function $f(x, y)$ is regular at every finite point of D_{pq}. The (p, q)-diagram D_{pq} is a region of the following character. Every analytic surface of the form*

$$(41) \quad x = \alpha z^p, \ y = \beta z^q, \qquad (|\alpha| + |\beta| \neq 0, \qquad z \text{ a complex parameter}),$$

cuts D_{pq} in a two-dimensional region $D_{pq}^{\alpha\beta}$. As the point $(x = \alpha/z^p, y = \beta/z^q)$ runs over $D_{pq}^{\alpha\beta}$ the points z run over the exterior of a closed convex region $\delta_{pq}^{\alpha\beta}$ in the z-plane. The convex region $\delta_{pq}^{\alpha\beta}$ is identical with the set of points $z = \rho e^{i\varphi}$ for which

$$(42) \qquad \rho \cos(\theta + \varphi) \leqq h(\alpha e^{ip\theta}, \beta e^{iq\theta}; p, q)$$

for every real value of θ. If z_0 is an extreme point[25] *of $\delta_{pq}^{\alpha\beta}$ other than the origin, then the point $(x = \alpha/z_0^p, y = \beta/z_0^q)$ is a singularity of the function $f(x, y)$.*

Let (x_0, y_0) be any finite point of D_{pq}. Then, by Definition 8 there exists a real value θ_0 such that (33) holds. Applying Lemma 2 with $\alpha = x_0$, $\beta = y_0$, $z = 1$, we see that $f(x, y)$ is regular at (x_0, y_0).

From its definition it is easily seen that D_{pq} is a region. We omit the details.

By property 16° the set of points described by (42) is a closed convex region. Let us denote this region by δ'. Let (x_0, y_0) be any (finite or infinite) point in $D_{pq}^{\alpha\beta}$ and write $x_0 = \alpha/z_0^p$, $y_0 = \beta/z_0^q$. Then, applying the definition of the diagram and property 14°, we see at once that the points z_0 are exterior to δ'. Conversely, if z_0 is any point exterior to δ', we easily see that the point $(x_0 = \alpha/z_0^p$, $y_0 = \beta/z_0^q)$ is in $D_{pq}^{\alpha\beta}$. Thus the region $\delta_{pq}^{\alpha\beta}$ described in the theorem is identical with δ'.

Let (α_0, β_0) be any pair of complex numbers such that $|\alpha_0| + |\beta_0| \neq 0$. Then from Lemma 3 and the relation of supporting lines to extreme points described in footnote 20, we see by an argument similar to one made by Pólya (*loc. cit.*,[7] p. 578) that if z_0 is an extreme point of $\delta_{pq}^{\alpha_0\beta_0}$ then the point $(\alpha = \alpha_0$,

[25] For the definition of an extreme point see footnote 19.

$\beta = \beta_0$, $z = z_0$) is a singular point of the function $g(\alpha, \beta, z) = (1/z)f(\alpha/z^p, \beta/z^q)$. If z_0 is different from zero this implies that the point $(x = \alpha_0/z_0^p,\ y = \beta_0/z_0^q)$ is a singular point of $f(x, y)$.

This concludes the proof of Theorem 3. Let us remark that in view of 15° the (p, q)-diagram D_{pq} certainly contains the complete Cartan (p, q)-region G_{pq} of Theorem 1. Thus D_{pq} furnishes an extended region of regularity of $f(x, y)$.

We will conclude this section with a few general remarks. By a consideration of the (p, q)-indicator $h(x, y; p, q)$ information concerning the question as to whether or not a particular point on the boundary of the complete Cartan (p, q)-region G_{pq} is a singular point of $f(x, y)$ may be obtained. Furthermore, from the manner of defining the functions $H(x, y; p, q)$, $h(x, y; p, q)$, it turns out that each of the functions $\log H(x, y; p, q)$, $h(x, y; p, q)$ is subharmonic whenever $H(x, y; p, q)$ is continuous. We shall not discuss any of these matters further.

9. **Functions of exponential type.** This section is devoted to the investigation of the special case $p = q = 1$. Throughout this section we write simply G, D, $h(x, y)$, etc. for G_{11}, D_{11}, $h(x, y; 1, 1)$, etc. For this special case various additional results, as well as more precision in the results already derived, may be obtained. For example, it is easily seen that the diagram D possesses an invariant property analogous to that possessed by the circular region G. Moreover it is possible to obtain information concerning the behavior of $f(x, y)$ at the infinite points by a consideration of the indicator $h(x, y)$. These aspects of the theory we do not consider here. However we do consider the question as to when a given region is the diagram of a function.

We say that a function

$$F(x, y) = \sum_{m,n=0}^{\infty} \frac{a_{mn} x^m y^n}{(m+n)!} \tag{43}$$

is of *exponential type* if the associated series (1) converges in a neighborhood of the origin. Thus the Borel transform $F(x, y)$ of a function $f(x, y)$ is a function of exponential type. We shall give a few examples of functions of exponential type and give without proof their associated functions and regions.

(i) $F(x, y) = e^{ax+by}$, $\qquad f(x, y) = 1/(1 - ax - by)$;

$H(x, y) = | ax + by |$, $\qquad G : | ax + by | < 1$;

$h(x, y) = R\{ax + by\}$, $\qquad D : ax + by \neq 1$.

(ii) $F(x, y) = x^j y^k e^{ax+by}$, $\qquad f(x, y) = (j + k)!\, x^j y^k/(1 - ax - by)^{j+k+1}$,

where j and k are positive integers. $H(x, y)$, G, $h(x, y)$ and D are the same as in (i).

(iii) Let $c_1, \cdots, c_r$ be any non-vanishing complex constants and let $a_1, \cdots, a_r$, $b_1, \cdots, b_r$ be any complex constants such that

$$(44)\qquad | a_j - a_k | + | b_j - b_k | \neq 0, \qquad j \neq k, \qquad j, k = 1, \cdots, r.$$

Then if

$$(45)\qquad F(x, y) = \sum_{j=1}^{r} c_j e^{a_j x + b_j y},$$

we have

$$f(x, y) = \sum_{j=1}^{r} \frac{c_j}{1 - a_j x - b_j y},$$

$$H(x, y) = \max_{j=1,\cdots,r} | a_j x + b_j y |, \qquad h(x, y) = \max_{j=1,\cdots,r} R\{a_j x + b_j y\}.$$

The region G consists of the product of all the regions $| a_j x + b_j y | < 1$, $j = 1, \cdots, r$. The diagram D is such that the convex region $\delta^{\alpha\beta}$ of Theorem 3 is the smallest convex region containing the points $a_j \alpha + b_j \beta$, $j = 1, \cdots, r$.

For the particular class of functions of the form (45) the circular region G and the diagram D possess an important geometrical property; namely, each is *planar convex*.[26] The fact that G is planar convex in this case is an immediate consequence of the fact that it is the intersection of the regions $| a_j x + b_j y | < 1$, each one of which is planar convex. The proof of the fact that the diagram D is planar convex involves the consideration of the convex regions $\delta^{\alpha\beta}$ of Theorem 3 and is somewhat more involved. We omit the details of the proof.

We shall investigate more closely the relation between the property of being a planar convex region and the property of being the diagram of a function of exponential type. We first prove the following theorem:

THEOREM 4. *Let C be any region possessing the following three properties:* (i) *it contains the origin;* (ii) *it is planar convex and* (iii) *every analytic plane through the origin*

$$(46)\qquad x = \alpha z, \quad y = \beta z, \quad (| \alpha | + | \beta | \neq 0, \quad z \text{ a complex parameter}),$$

cuts it in a (plane) region $C^{\alpha\beta}$ which is such that as the point (x, y) runs over $C^{\alpha\beta}$ the "inverted" point

$$(47)\qquad x' = \alpha^2/x, \qquad y' = \beta^2/y,$$

runs in the analytic plane (46) *over the exterior of a closed convex (plane) region $J^{\alpha\beta}$. Then the region C is the diagram of a function of exponential type.*[27]

[26] Planar convexity is a concept due to H. Behnke and E. Peschl, "Konvexität in bezug auf analytische Ebenen im kleinen und grossen," *Mathematische Annalen*, 111 (1935), pp. 158–177. For *schlicht* regions planar convexity is defined as follows: A *schlicht* region R is said to be *planar convex* if through every boundary point of R there passes an analytic plane which never cuts the interior of R.

[27] For the sake of completeness let us remark that the diagram of every function of exponential type possesses the geometrical property expressed by (iii); indeed, this property is essentially a restatement of the property of the diagram named in Theorem 3 for the case $p = q = 1$.

For the proof of Theorem 4 let us consider a point-set A consisting of every point (x, y) which is a boundary point of any of the regions $C^{\alpha\beta}$. Let (α_ν, β_ν), $\nu = 0, 1, 2, \cdots$, be an everywhere dense set of points of A. Then since C is planar convex and since each of the points (α_ν, β_ν) is a boundary point of C it follows from the definition of planar convexity that through each of the points (α_ν, β_ν) there passes an analytic plane which never cuts C. Since C contains the origin the constant term in the equation of any such plane can not vanish. From these facts it is clear that there exists a finite or infinite sequence of distinct analytic planes

$$a_s x + b_s y = 1, \qquad s = 0, 1, 2, \cdots, \tag{48}$$

such that each of the points (α_ν, β_ν), $\nu = 0, 1, 2, \cdots$, is on one and only one such plane and such that none of the planes cuts C. Notationally we proceed as though the sequence (48) were infinite.

Let c_s, $s = 0, 1, 2, \cdots$, be a sequence of non-vanishing complex constants such that the series

$$\sum_{s=0}^{\infty} \frac{c_s}{[|a_s| + |b_s|]^s} \tag{49}$$

is absolutely convergent and let us consider the series

$$f(x, y) = \sum_{s=0}^{\infty} \frac{c_s}{1 - a_s x - b_s y}. \tag{50}$$

Series of the form $\sum_{s=0}^{\infty} \frac{\gamma_s}{z - \delta_s}$ have been investigated by various writers.[28] We shall not carry through a detailed discussion of the properties of the series (50). Let us merely state a few facts concerning the series. The series (50) represents an analytic function $f(x, y)$ regular in C. Using the geometrical property of C expressed in Theorem 4 by (iii), we may show that the function $f(x, y)$ defined by (50) has a singularity at each of the points (α_ν, β_ν), $\nu = 0, 1, 2, \cdots$. Using these facts and the property (iii) again, we may show that the Borel transform $F(x, y)$ of $f(x, y)$ has C as its diagram D. We shall give no more of the details of the proof of the theorem. In the interest of completeness let us remark that the Borel transform of (50) has the form

$$F(x, y) = \sum_{s=0}^{\infty} c_s e^{a_s x + b_s y}.$$

The question arises: Is the diagram D of every function of exponential type planar convex? The answer is in the negative; that is, *there exists a function $F(x, y)$ of exponential type whose diagram D is not planar convex.* Behnke has

[28] See, for example; A. Pringsheim, "Über bemerkenswerte Singularitätenbildungen bei gewissen Partialbruchreihen," *Sitzungsbericht der Bayerische Akademie*, (1927), pp. 145–164, where also various references are given.

shown that there exists a function $\varphi(x, y)$ which is regular at every point of the complete circular region

$$|x^2 + y^2| < 1, \tag{51}$$

which is not continuable over its boundary and which has the circular region (51) as its invariant convergence region.[29] In view of Theorem 1 and equation (14) we see that the Borel transform $\Phi(x, y)$ of $\varphi(x, y)$ has the type-function $H(x, y) = |x^2 + y^2|^{1/2}$ and associated region G the region defined by (51). Furthermore, since $\varphi(x, y)$ is not continuable over the boundary of G, the diagram D of $\Phi(x, y)$ is identical with the circular region G and hence the indicator has the form $h(x, y) = |x^2 + y^2|^{1/2}$. By noting, for example, that the point (1, 0) is on the boundary of D and that every analytic plane through this point cuts D, we see that D is not planar convex, a fact which may also be seen by noting that the equation of the boundary $|x^2 + y^2| = 1$ does not satisfy the necessary conditions for planar convexity given by Behnke and Peschl.[30]

III. Summability. Star Regions

10. **Summability by Borel's integral means.** In this section we investigate the summability of series of the form (3) by Borel's integral means and relate the results to the problem of analytic continuation of the sum-function $f(x, y)$. We first recall certain definitions and results in the theory of summation of series.[31]

A series

$$\sum_{\nu=0}^{\infty} \alpha_\nu \tag{52}$$

is said to be *absolutely summable* (*by Borel's integral means*) if each of the integrals

$$\int_0^\infty e^{-a} F(a)da, \qquad \int_0^\infty e^{-a} |F^{(\lambda)}(a)| da, \qquad \lambda = 0, 1, 2, \cdots,$$

exists, where

$$F(a) = \sum_{\nu=0}^{\infty} \frac{\alpha_\nu a^\nu}{\nu!}.$$

The first of these integrals is called the *sum* of the series (52)

18°. *An absolutely convergent series is absolutely summable.*

19°. *If a series*

$$p(z) = \sum_{\nu=0}^{\infty} \alpha_\nu z^\nu \tag{53}$$

---

[29] See Behnke, *loc. cit.* 15, especially Theorems 11 and 12. The region in (51) is a particular region of a general class of regions possessing these properties.

[30] *Loc. cit.* 26. See Theorem 7a.

[31] See, for example, T. Fort, *Infinite series*, Oxford (1930), pp. 222–233. See also the references in footnote 4.

is absolutely summable at a point M then it is absolutely summable for every point z on the straight line segment OM. Furthermore the function p(z) is an analytic function regular interior to and on the circumference of the circle having OM as a diameter. The analytic continuation of p(z) in this circle is represented by the integral

$$\int_0^\infty e^{-a}P(az)da, \qquad \left(P(z) = \sum_{\nu=0}^{\infty} \frac{\alpha_\nu z^\nu}{\nu!}\right).$$

20°. *If a line is drawn from the origin to each singular point of p(z) and that portion of the plane bounded by perpendiculars to these lines at the singular points and including the origin is denoted by b; then* (53) *is absolutely summable within b to a function analytic at every point within b. It is not absolutely summable at any point outside b.*

This concludes the resume of results from the classical theory. We next prove the following theorem:

THEOREM 5. *Let* (1) *be any series convergent in a neighborhood of the origin and let us consider the series* (3) *associated with* (1). *If* (3) *is absolutely summable (by Borel's integral means) for every value of* (x, y) *in a neighborhood of a point* (x_0, y_0) *then* (3) *is absolutely summable in a neighborhood of every point of the curve-segment*

$$x = x_0\rho^p, \qquad y = y_0\rho^q, \qquad 0 \leqq \rho \leqq 1. \tag{54}$$

Furthermore the function $f(x, y)$ *is regular at every point* (x, y) *which is expressible in the form*

$$x = x_0 z^p, \qquad y = y_0 z^q, \qquad | z - \tfrac{1}{2} | \leqq \tfrac{1}{2}. \tag{55}$$

The analytic continuation of $f(x, y)$ *in this set of points is represented by the integral*

$$\int_0^\infty e^{-a} F(xa^p, ya^q; p, q)da, \tag{56}$$

where $F(x, y; p, q)$ *is the entire function defined by the series* (2).

For the proof let δ be a positive number such that the series (3) is absolutely summable for $| x - x_0 | < \delta$, $| y - y_0 | < \delta$. Then each of the series

$$p(\alpha, \beta, z) = f(\alpha z^p, \beta z^q) = \sum_{\nu=0}^{\infty} z^\nu \left\{ \sum_{pm+qn=\nu} a_{mn}\alpha^m\beta^n \right\}, \tag{57}$$

where $| \alpha - x_0 | < \delta$, $| \beta - y_0 | < \delta$, is absolutely summable at $z = 1$ and hence by 19° each is absolutely summable for z on the straight line segment joining the points $z = 0$, $z = 1$. From these facts we see at once that (3) is absolutely summable in the neighborhood

$$| x - x_0\rho_0^p | < \rho_0^p\,\delta, \qquad | y - y_0\rho_0^q | < \rho_0^q\,\delta,$$

of any point $(x_0\rho_0^p, y_0\rho_0^q)$ for which $0 < \rho_0 \leqq 1$. Furthermore, since (1) is absolutely convergent in a neighborhood of the origin, (3) is absolutely con-

vergent in a neighborhood of the origin and hence by 18° the series (3) is absolutely summable in a neighborhood of the origin. These results prove that (3) is absolutely summable in a neighborhood of every point of the curve-segment (54).

That $f(x, y)$ is regular at every point (x, y) expressible in the form (55) we see as follows. Retaining the notation of the previous paragraph, we see by 19° that each of the functions $p(\alpha, \beta, z)$, where $|\alpha - x_0| < \delta$, $|\beta - y_0| < \delta$, as a function of z alone, is regular for z in the circle $|z - \frac{1}{2}| \leqq \frac{1}{2}$. Moreover, since $f(x, y)$ is regular at the origin, it follows that $p(\alpha, \beta, z)$, as a function of α, β, z, is regular for $|\alpha - x_0| < \delta$, $|\beta - y_0| < \delta$ for $|z|$ sufficiently small. Applying repeatedly a classical result due to Hartogs,[32] we easily obtain the result that the function $p(\alpha, \beta, z)$, as a function of α, β, z, is regular for $|\alpha - x_0| < \delta$, $|\beta - y_0| < \delta$, $|z - \frac{1}{2}| \leqq \frac{1}{2}$. From this we see that the function $f(x, y)$ is surely regular at every point expressible in the form (55).

From 19° it follows that the integral (56) represents the analytic continuation of $f(x, y)$ in the set of points indicated.

This concludes the proof of the theorem. We next define a four-dimensional analogue of the Borel polygon of summability and investigate its properties.

Let $f(x, y)$ be any function regular at the origin. Let (α, β) be any pair of complex numbers such that $|\alpha| + |\beta| \neq 0$. In the z-plane draw a line from the origin to each point z' which is such that the point $x = \alpha z'^p$, $y = \beta z'^q$ is a singular point of $f(x, y)$. Denote that (open) portion of the z-plane bounded by perpendiculars to these lines at the points z' and including the origin by $b_{pq}^{\alpha\beta}$. As the point z runs over $b_{pq}^{\alpha\beta}$ let the point $(x = \alpha z^p, y = \beta z^q)$ run over a point-set $B_{pq}^{\alpha\beta}$. Denote by B_{pq} the set of points in the x, y-space consisting of every point (x, y) which is in any $B_{pq}^{\alpha\beta}$, $(|\alpha| + |\beta| \neq 0)$. We shall call B_{pq} the *Borel (p, q)-region of summability* of $f(x, y)$.

From its construction it follows that B_{pq} is a region. The theorem which we next prove relates B_{pq} to the region of summability of (3).

THEOREM 6. *Let $f(x, y)$ be any function regular at the origin and let B_{pq} be its Borel (p, q)-region of summability. Then the series development* (3) *of $f(x, y)$ is absolutely summable (by Borel's integral means) in a neighborhood of every point of B_{pq}. It is not absolutely summable in a neighborhood of any point exterior to B_{pq}.*

Let (x_0, y_0) be any point in B_{pq} (other than the origin). Recalling the construction of B_{pq}, we see that there exists a positive constant δ such that $f(x, y)$ is regular at every point (x, y) which is expressible in the form

$$(58)\quad x = \alpha z^p, y = \beta z^q, \qquad |z - \tfrac{1}{2}| \leqq \tfrac{1}{2}, |\alpha - x_0| < \delta, |\beta - y_0| < \delta.$$

By 20° this implies that each of the series (57) with $|\alpha - x_0| < \delta$, $|\beta - y_0| < \delta$, is absolutely summable for $z = 1$ which is equivalent to the fact that (3) is absolutely summable for every point for which $|x - x_0| < \delta$, $|y - y_0| < \delta$.

[32] See footnotes 24, 12.

As we remarked in the proof of Theorem 5 the series (3) is absolutely summable in a neighborhood of the origin. These two results show that (3) is absolutely summable in a neighborhood of every point of B_{pq}.

Next let (x_0, y_0) be any point in a neighborhood of which (3) is absolutely summable. Then by Theorem 5 we see that $f(x, y)$ is regular at every point (x, y) expressible in the form (55). By the construction of $b_{pq}^{x_0y_0}$ this implies that the point $z = 1$ is in $b_{pq}^{x_0y_0}$ and hence the point (x_0, y_0) is in B_{pq}.

This concludes the proof of Theorem 6.

Let us associate with every point (x_0, y_0) in B_{pq} the set of points (x, y) expressible in the form (55) and let us denote by $\mathfrak{B}_{pq}$ the totality of all points (x, y) obtained in this manner. It is evident that $\mathfrak{B}_{pq}$ is a region. Furthermore, from Theorems 5 and 6 we see that $f(x, y)$ is regular throughout the interior of $\mathfrak{B}_{pq}$ and that the analytic continuation of $f(x, y)$ in $\mathfrak{B}_{pq}$ is represented by the integral (56). We shall call $\mathfrak{B}_{pq}$ the *Borel* (p, q)*-region of regularity* of $f(x, y)$.

We shall state without proofs some of the relations existing among the various regions which we have studied so far. They are easily seen to be related as follows:

$$G_{pq} \subseteq B_{pq} \subseteq \mathfrak{B}_{pq} \subseteq D_{pq}.$$

Furthermore if the indicator $h(x, y; p, q)$ is everywhere non-negative then $\mathfrak{B}_{pq} \equiv D_{pq}$. If, however, $h(x, y; p, q)$ is sometimes negative then D_{pq} contains points which $\mathfrak{B}_{pq}$ does not contain and thus in this case the method of analytic continuation by the use of the functional transform $F(x, y; p, q)$ furnishes more information than does the method using Borel's theory of summability.

11. **Star regions.** Let $f(x, y)$ be any function regular at the origin. Let σ and μ be any two positive numbers. Consider a curve

$$x = \alpha\rho^\sigma, \qquad y = \beta\rho^\mu, \qquad 0 \leqq \rho < \infty, \tag{59}$$

where α, β are any complex numbers such that $|\alpha| + |\beta| \neq 0$. If the analytic continuation of $f(x, y)$ has no singular point along this curve, we retain the whole curve. If $x = \alpha\rho_0^\sigma$, $y = \beta\rho_0^\mu$ is the first singular point of $f(x, y)$ along this curve, we retain only the curve up to the first singular point excluding that point. With the same construction for every pair of complex numbers for which $|\alpha| + |\beta| \neq 0$, we obtain a portion of the x, y-space (the infinite points omitted) which we call the (σ, μ)*-star region* of $f(x, y)$ and which we denote by $S_{\sigma\mu}$. If $x_0 = \alpha\rho_0^\sigma$, $y_0 = \beta\rho_0^\mu$ is the first singular point of $f(x, y)$ on the curve (59) then we call the point (x_0, y_0) the *vertex of the* (σ, μ)*-star region belonging to the curve* (59).

Let a be any positive number and let us define the function

$$F_a(x, y; \sigma, \mu) = \sum_{m,n=0}^{\infty} \frac{a_{mn} x^m y^n}{\Gamma(1 + a\sigma m + a\mu n)}, \tag{60}$$

associated with the power-series development (1) of the function $f(x, y)$. The function $F_a(x, y; \sigma, \mu)$ is obviously an entire function. Let us define two real-valued functions

$$(61)\qquad c^*(x, y; \sigma, \mu) = \lim_{a=0} \limsup_{\omega=\infty} \{a \log \log^+ | F(xe^{\sigma\omega}, ye^{\mu\omega}; \sigma, \mu) | - \omega\},$$

$$(62)\qquad c(x, y; \sigma, \mu) = \limsup_{x'=x, y'=y} c^*(x', y'; \sigma, \mu),$$

where $\log^+ A$ denotes $\log A$ if $A \geqq 1$, and 0 if $A < 1$.

The main result of this section is the following:

THEOREM 7. *The (σ, μ)-star region $S_{\sigma\mu}$ of a function $f(x, y)$ regular at the origin is identical with the set of points (x, y) for which $c(x, y; \sigma, \mu) < 0$ and its vertices are given by the points (x, y) for which $c(x, y; \sigma, \mu) = 0$.*

From the theory of Dirichlet series[33] we have the following results:

21°. (i) *The limit in* (61) *exists whenever $f(x, y)$ is any function regular at the origin.* (ii) *The function $f(xe^{-\sigma s}, ye^{-\mu s})$, as a function of the complex variable s for fixed (x, y), is regular for s on the ray $I(s) = 0$, $R(s) > c^*(x, y; \sigma, \mu)$.* (iii) *The function $f(xe^{-\sigma s}, ye^{-\mu s})$, as a function of s, has a singularity at the point $s = c^*(x, y; \sigma, \mu)$.*

We proceed with the proof of Theorem 7. Denoting by $S'_{\sigma\mu}$ the set of points (x, y) for which $c(x, y; \sigma, \mu) < 0$, we shall first show that $S_{\sigma\mu} \subseteq S'_{\sigma\mu}$. Let $P(x_0, y_0)$ be any point in $S_{\sigma\mu}$. Then there exists a neighborhood U of P such that every point in U is also in $S_{\sigma\mu}$. In view of the character of $S_{\sigma\mu}$ it follows that $f(x, y)$ is regular at every point $(\rho^\sigma x', \rho^\mu y')$, where $0 \leqq \rho \leqq 1$ and (x', y') is any point in U. Hence $f(xe^{-\sigma s}, ye^{-\mu s})$, as a function of x, y and s, is surely regular for (x, y) in U and for s on the ray $I(s) = 0$, $R(s) > 0$. By 21° this implies that $c^*(x, y; \sigma, \mu) \leqq 0$ for (x, y) in U, which by (62) implies that $c(x, y; \sigma, \mu) \leqq 0$ for (x, y) in U. Using the easily proved property

$$(63)\qquad c(x\rho^\sigma, y\rho^\mu; \sigma, \mu) = c(x, y; \sigma, \mu) + \log \rho, \; 0 < \rho < \infty$$

and noting that there exists a value $\rho_0 > 1$ such that the point $(x_0\rho_0^\sigma, y_0\rho_0^\mu)$ is in U, we see that $c(x_0, y_0; \sigma, \mu) < 0$. Hence P is in $S'_{\sigma\mu}$ and consequently $S_{\sigma\mu} \subseteq S'_{\sigma\mu}$.

Next let $P(x_0, y_0)$ be any point in $S'_{\sigma\mu}$. Then there exist a positive number δ and a neighborhood U_δ of P such that

$$(64)\qquad c^*(x, y; \sigma, \mu) < c(x_0, y_0; \sigma, \mu) + \delta < 0$$

for (x, y) in U_δ. Hence by 21° each of the functions $f(xe^{-\sigma s}, ye^{-\mu s})$, as a function of s for (x, y) in U_δ, is regular for s on the ray $I(s) = 0$, $R(s) > c(x_0, y_0; \sigma, \mu) + \delta$. Furthermore, since $f(x, y)$ is regular at the origin, the function $f(xe^{-\sigma s}, ye^{-\mu s})$, as a function of x, y and s, is regular for (x, y) in U_δ for $R(s)$ sufficiently large.

[33] See M. Riesz, *loc. cit.* 6; V. Bernstein, *op. cit.* 6.

Applying Hartogs' theorem,[34] we obtain without difficulty the result that $f(xe^{-\sigma s}, ye^{-\mu s})$, as a function of x, y and s, is regular for (x, y) in U_δ and for s on the ray $I(s) = 0$, $R(s) > c(x_0, y_0; \sigma, \mu) + \delta$. Hence, using (64), we have the result that $f(xe^{-\sigma s}, ye^{-\mu s})$, as a function of x, y and s, is regular for (x, y) in U_δ and for $s = 0$, which implies that $f(x, y)$ is regular at P. From (63) we see that with (x_0, y_0) the set $S'_{\sigma\mu}$ contains also all the points $(x_0\rho^\sigma, y_0\rho^\mu)$ with $0 \leqq \rho \leqq 1$. This fact and the regularity of $f(x, y)$ at the point $P(x_0, y_0)$ prove that P is in $S_{\sigma\mu}$. Hence $S'_{\sigma\mu} \subseteq S_{\sigma\mu}$.

Combining the results of the two previous paragraphs, we have the result that $S'_{\sigma\mu} = S_{\sigma\mu}$, which proves the first part of the theorem. From 21° and (62) we see that every point (x_0, y_0) for which $c(x_0, y_0; \sigma, \mu) = 0$ is a singular point of $f(x, y)$ and indeed that it is the first one on the curve (59) formed with $(\alpha, \beta) = (x_0, y_0)$. If (x_0, y_0) is a vertex of $S_{\sigma\mu}$ then it is clear from (63) that $c(x_0, y_0; \sigma, \mu) = 0$.

This concludes the proof of Theorem 7.

PRINCETON UNIVERSITY AND THE INSTITUTE FOR ADVANCED STUDY.

---

[34] See footnotes 24, 12. The details here are entirely analogous to those in the proof given by Bochner and Martin (*loc. cit.*, footnote 24) of their theorem C relating the star vertices of a function of several variables to the vertices of certain associated functions of a single variable.

ANNALS OF MATHEMATICS
Vol. 38, No. 3, July, 1937

A NOTE ON COMPLETELY CONTINUOUS TRANSFORMATIONS

BY F. SMITHIES

(Received December 14, 1936)

1. INTRODUCTION

In this note we shall consider completely continuous transformations between the function spaces L^p and L^q, where $p > 1$, $q > 1$, and also between the corresponding spaces S^p and S^q, whose elements are infinite sequences of complex numbers.

We use the Lebesgue integral throughout, and suppose all functions introduced to be measurable.

A complex function $f(s)$ of the real variable s is said to belong to the space L^p, where $p > 1$, if

$$\int_{-\infty}^{\infty} |f(s)|^p \, ds < \infty.$$

We write $f \epsilon L^p$. The norm $\| f \|_p$ of f is defined by the equation

$$\| f \|_p = \left\{ \int_{-\infty}^{\infty} |f(s)|^p \, ds \right\}^{1/p}.$$

Similarly, an infinite sequence $\{x_n\}$ of complex numbers is said to belong to the space S^p, where $p > 1$, if

$$\sum_{n=1}^{\infty} | x_n |^p < \infty.$$

We write $x \epsilon S^p$, and define the norm $\| x \|_p$ of x by the equation

$$\| x \|_p = \left(\sum_{n=1}^{\infty} | x_n |^p \right)^{1/p}.$$

We consider transformations of the forms

$$(1) \qquad g(s) = \int_{-\infty}^{\infty} K(s, t) f(t) \, dt \qquad (-\infty < s < \infty)$$

and

$$(2) \qquad y_m = \sum_{n=1}^{\infty} a_{mn} x_n \qquad (m = 1, 2, \cdots),$$

where $f \epsilon L^p$, $g \epsilon L^q$, $x \epsilon S^p$, $y \epsilon S^q$. We shall obtain sufficient conditions that the transformations defined by (1) and (2) should be completely continuous, i.e. that they should transform every bounded set into a compact set.

Banach [1], pp. 98–99, proved that a sufficient condition for the transformation

$$g(s) = \int_0^1 K(s, t) f(t)\, dt \qquad (0 \leqq s \leqq 1) \tag{3}$$

to be completely continuous is that

$$\int_0^1 \int_0^1 | K(s, t) |^r \, ds\, dt < \infty, \tag{4}$$

where

$$r = \mathrm{Max}[p', q],$$

and p' is the conjugate index to p, i.e. $(1/p) + (1/p') = 1$. This result does not hold for an infinite interval.

Cohen [2] obtained a similar result for sequences, viz. that the condition

$$\sum_{m,n=1}^{\infty} | a_{mn} |^r < \infty, \tag{5}$$

where

$$r = \mathrm{Min}[p', q],$$

is sufficient for (2) to define a completely continuous transformation between S^p and S^q.

For the special case $p = q$, conditions have been given by Hille and Tamarkin [3] for the transformation (1), and by Cohen [2] for (2), which are identical with those given in Theorems 1 and 2 below.

The theorems proved in this paper include all the above results, and, as we shall show in §3, go somewhat further.

2. General Theorems

Let us define, for given $p(>1)$ and $q(>1)$,

$$|| K || = \left\{ \int_{\infty}^{\infty} ds \left(\int_{-\infty}^{\infty} | K(s, t) |^{p'} dt \right)^{q/p'} \right\}^{1/q},$$

where $(1/p) + (1/p') = 1$.

Theorem 1. *If*

$$\| K \| < \infty, \tag{6}$$

then the equation

$$g(s) = \int_{-\infty}^{\infty} K(s, t) f(t)\, dt \tag{7}$$

defines a completely continuous transformation from L^p into L^q.

We require the following result.

Lemma. *If $\| K \| < \infty$, then, given $\epsilon(> 0)$, we can find a real number N and a function $G(s, t)$ such that*

(i) $$G(s, t) = 0 \qquad (| s | > N \text{ or } | t | > N),$$

(ii) *$G(s, t)$ is continuous in $-N \leqq s \leqq N$, $-N \leqq t \leqq N$, and*

(iii) $$\| K - G \| < \epsilon.$$

We omit the proof, which is of a familiar type. For a similar argument, see Hobson [4], p. 250.

By the lemma, there is a sequence of functions $K_n(s, t)$, each continuous in a certain square and vanishing identically outside it, such that

$$\| K - K_n \| \to 0.$$

For each integer n, the equation

$$g(s) = \int_{-\infty}^{\infty} K_n(s, t) f(t)\, dt$$

defines a completely continuous transformation from L^p into L^q. See Banach [1], p. 97. The fact that the interval of definition of $g(s)$ is apparently infinite presents no difficulty, for $g(s)$ vanishes outside a finite interval depending only on n.

If
$$g(s) = \int_{-\infty}^{\infty} K(s, t) f(t)\, dt,$$

we see by Hölder's inequality that $\| g \|_q \leqq \| K \| \cdot \| f \|_p$. Let $| K |$ be the norm of the transformation determined by $K(s, t)$, as defined by Banach [1], p. 54. Then $| K |$ is the least number such that $\| g \|_q \leqq | K | \cdot \| f \|_p$ for all f in L^p. Hence we must have $| K | \leqq \| K \|$.

Since $\| K - K_n \| \to 0$, we must have $| K - K_n | \to 0$. By Banach [1], p. 96, Theorem 2, the limit of a sequence of completely continuous transformations convergent according to the norm is completely continuous. This is the required result.

THEOREM 2. *Consider the transformation*

$$y_m = \sum_{n=1}^{\infty} a_{mn} x_n, \tag{8}$$

where $x \in S^p$. Then a sufficient condition that it be a completely continuous transformation from S^p into S^q is that

$$\sum_{m=1}^{\infty} \left(\sum_{n=1}^{\infty} | a_{mn} |^{p'} \right)^{q/p'} < \infty. \tag{9}$$

We shall show that this is a consequence of Theorem 1.

Define

$$\begin{aligned} K(s, t) &= 0 \quad (s < 0 \text{ or } t < 0), \\ &= a_{mn} \ (m - 1 \leqq s < m, n - 1 \leqq t < n), \end{aligned}$$

$$f(t) = x_n \ (n - 1 \leqq t < n). \tag{10}$$

Then, when $m - 1 \leqq s < m$,

$$g(s) = \int_{-\infty}^{\infty} K(s, t) f(t)\, dt \tag{11}$$

$$= \sum_{n=1}^{\infty} a_{mn} x_n$$

$$= y_m .$$

Furthermore $$\| f \|_p = \| x \|_p ,$$

and $$\| g \|_q = \| y \|_q .$$

Let T be any bounded set of elements x in S^p, so that, for some A, $\| x \|_p \leqq A$ for all x in T. Let T_1 be the set of functions $f(t)$ of L^p corresponding to the elements x in T by (10). Let T_2 be the set of functions $g(s)$ of L^q obtained from the functions $f(t)$ of T_1 by equation (11), and let T_3 be the corresponding set of elements y of S^q.

By Theorem 1, T_2 is compact in L^q. Hence any infinite subset of T_2 contains a subsequence $\{g^{(\nu)}(s)\}$ convergent in L^q. It follows that the corresponding sequence in T_3 is convergent in S^q, i.e. T_3 is compact. This proves Theorem 2.

We now state two further theorems, which can be deduced from Theorem 1 in the same way as Theorem 2.

THEOREM 3. *Consider the transformation given by*

$$g(s) = \sum_{n=1}^{\infty} k_n(s) x_n . \tag{12}$$

If

$$\int_{-\infty}^{\infty} \left\{ \sum_{n=1}^{\infty} | k_n(s) |^{p'} \right\}^{q/p'} ds < \infty , \tag{13}$$

then (12) *defines a completely continuous transformation from S^p into L^q.*

THEOREM 4. *Consider the transformation given by*

$$y_m = \int_{-\infty}^{\infty} k_m(t)\, f(t)\, dt. \tag{14}$$

If

$$\sum_{m=1}^{\infty} \left\{ \int_{-\infty}^{\infty} | k_m(t) |^{p'} dt \right\}^{q/p'} < \infty , \tag{15}$$

then (14) *defines a completely continuous transformation from L^p into S^q.*

3. EXAMPLES

We now discuss a few examples indicating the scope of the above theorems.

(i) Let
$$K(s, t) = s^{-\alpha} t^{-\beta} \quad \text{in } 0 < s \leqq 1,\ 0 < t \leqq 1,$$
$$= 0 \quad \text{elsewhere.}$$

Then (6) holds if and only if $\beta p' < 1$, $\alpha q < 1$. Condition (4) is equivalent to $\alpha r < 1$, $\beta r < 1$, where

$$r = \mathrm{Max}[p', q].$$

Thus (6) is more inclusive than (4).

(ii) Let $$a_{mn} = m^{-\alpha} n^{-\beta}.$$

Then (9) holds if and only if $\beta p' > 1$, $\alpha q > 1$. Condition (5) is equivalent to $\alpha r > 1$, $\beta r > 1$, where

$$r = \mathrm{Min}[p', q].$$

Thus (9) is more inclusive than (5).

(iii) Each of the conditions (6), (9), (13), and (15) is best possible of its kind. E.g. consider (6).

Take
$$\begin{aligned} K(s, t) &= s^{-\alpha} t^{-\beta} && \text{in } 0 < s \leqq \tfrac{1}{2},\ 0 < t \leqq \tfrac{1}{2}, \\ &= 0 && \text{elsewhere.} \end{aligned}$$

(6) is equivalent to $\beta p' < 1$, $\alpha q < 1$. Suppose that $\beta p' \geqq 1$, say $\beta p' = 1$. Let
$$f(t) = t^{-1/p} (\log t)^{-\gamma},$$
where $1/p < \gamma < 1$. Then $f \in L^p$, but

$$\int_0^{\frac{1}{2}} K(s, t) f(t)\, dt$$

does not exist for any value of s in $(0, \frac{1}{2})$.

(iv) Now suppose that $\beta p' < 1$, $\alpha q = 1$. Then, for any f in L^p,

$$\begin{aligned} g(s) &= \int_0^{\frac{1}{2}} K(s, t) f(t)\, dt \\ &= A s^{-1/q}, \end{aligned}$$

so that $g(s)$ does not belong to L^q.

Similar examples can easily be given for conditions (9), (13), and (15).

REFERENCES

1. S. Banach: *Théorie des opérations linéaires* (Warszawa, 1932).
2. L. W. Cohen: "Transformations on spaces of infinitely many dimensions." *Ann. Math.* (2) 37 (1936), 326–335.
3. E. Hille, J. D. Tamarkin: "On the theory of linear integral equations. II." *Ann. Math.* (2) 35 (1934), 445–455.
4. E. W. Hobson: *Theory of functions of a real variable*, vol. 2 (Cambridge, 1926).

Princeton, N. J.

ANNALS OF MATHEMATICS
Vol. 38, No. 3, July, 1937

PEANIAN CONTINUA NOT IMBEDDABLE IN A SPHERICAL SURFACE[1]

BY SCHIEFFELIN CLAYTOR

(Received December 14, 1936)

INTRODUCTION

This paper is devoted to the demonstration of the following

THEOREM. *A peanian continuum which is not homeomorphic with a subset of the surface of a sphere necessarily contains either a primitive skew curve*[2] *or a topological image of one of the following figures*:

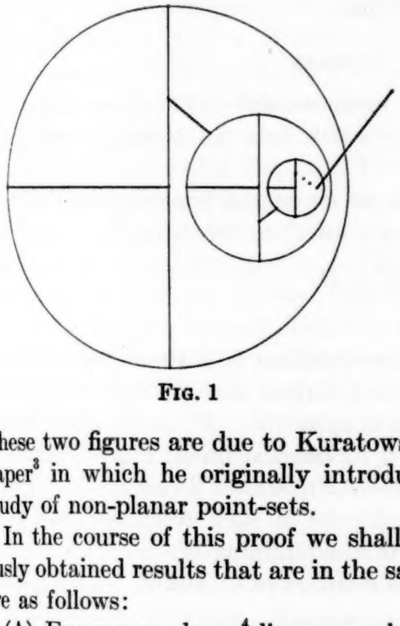


FIG. 1

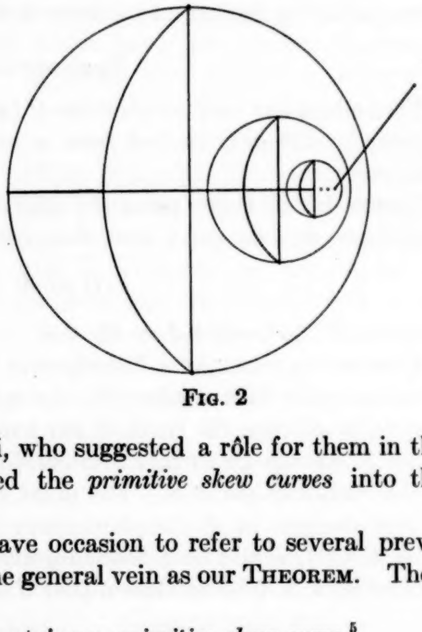


FIG. 2

These two figures are due to Kuratowski, who suggested a rôle for them in the paper[3] in which he originally introduced the *primitive skew curves* into the study of non-planar point-sets.

In the course of this proof we shall have occasion to refer to several previously obtained results that are in the same general vein as our THEOREM. They are as follows:

(A) Every non-planar[4] linear graph contains a *primitive skew curve*.[5]

---

[1] Presented to the American Mathematical Society, December 30, 1936.

[2] For the definition of a *primitive skew curve* consult the Introduction to the author's *Topological Immersion of Peanian Continua in a Spherical Surface*, Annals of Mathematics, Vol. 35, No. 4, (1934), pp. 809–835. This paper will hereafter be referred to as T.I.

[3] Kuratowski, *Sur les problèmes des courbes gauches en Topologie*, Fund. Math., XV, pp. 271–283. See page 272, footnote 2, for a description of the construction of figures 1 and 2.

[4] A *planar* point set is one which is homeomorphic with some subset of the plane.

[5] Cf. Kuratowski, loc. cit.; and H. Whitney, *Non-separable and planar graphs*, Trans. Amer. Math. Soc., vol. 34, (1932), pp. 339–362.

(B) If K is a peanian continuum which is not homeomorphic with a subset of a spherical surface, then either

1) K contains a primitive skew curve, or if not, then

2) K has a cut-point P which is not a boundary point of the closure of *every* component of $K - P$.

Special case of (B):

(C) Every *cyclic* peanian continuum which is not homeomorphic with a subset of a spherical surface contains a *primitive skew curve.*[6]

It will be seen at a glance that our result (C) compares favorably with the original result (A) of Kuratowski in the matter of the simplicity of the ideas involved. But in passing to the treatment of the general case (B) the author was obliged in T. I. to introduce the somewhat cumbersome notion of *boundary point.* We have managed here, however, to employ the remaining (and as yet unused) pair of Kuratowski's curves in altogether freeing our results of auxiliary ideas involving the use of intricate definitions.

Preliminary Lemmas

Throughout our work we shall use C (and upon occasion C_n) to represent a cyclic peanian continuum, distinct from a point, which does not contain a primitive skew curve.

Lemma 1. *If P is a point of C that does not lie on any boundary curve*[7] *of this continuum, then for any $\epsilon > 0$, there exists a normal representation,*[8]

$$C = A + J + B,$$

such that $B \supset P$, and $\delta(J + B) < \epsilon$.

Demonstration: As a consequence of proposition (C)[9] the continuum C is homeomorphic with a subset C_1 of a spherical surface S_1. If it should happen that $C_1 = S_1$ then the truth of our lemma is apparent. If, on the other hand, $C_1 \neq S_1$ we can go further and consider a topological image C', $= t(C)$, of C in the euclidean plane S. The point P', $= t(P)$, cannot lie on the frontier F' of any domain in S complementary to C'; for if such were the case then F, $= t^{-1}(F')$, would be a *boundary curve* of C containing P.

Now let $\delta > 0$ be so chosen that if X is a subset of C, and $X' = t(X)$, then

$$(1) \qquad \delta(X') < \delta \quad \text{implies} \quad \delta(X) < \epsilon.$$

According to a theorem of Whyburn[10] there exists in C' a simple closed curve J' of diameter $<\delta$ that encloses the point P'. Let A' be the subset of C'

[6] Cf. Introduction to T.I.

[7] See Definition I of T.I.

[8] Consult page 813 of T.I. for the definition of a *normal representation.*

[9] Cf. Introduction.

[10] *Concerning continua in the plane,* Trans. Amer. Math. Soc. Vol. 29, No. 2, pp. 369–400. Theorem 7.

that lies *outside* of J', and let B' denote the subset of C' that lies *within* J'. The curve J' evidently *separates* A' and B' in C', and furthermore constitutes a *boundary curve* of each of the two continua $A' + J'$ and $J' + B'$. Consequently $A' + J' + B'$ is a *normal representation* of C', and on setting $A = t^{-1}(A')$, $J = t^{-1}(J')$, $B = t^{-1}(B')$ we get the desired normal representation $C = A + J + B$; for, $B' \supset P'$ implies $B \supset P$ and, in view of (1), the inequality $\delta(J' + B') < \delta$ insures $\delta(J + B) < \epsilon$.

LEMMA 2. *In a normal representation $C = A + J + B$, let P be a point $\subset B$ that does not lie on a boundary curve of C; then neither will P lie on a boundary curve of $J + B$.*

DEMONSTRATION: Suppose on the contrary that P does lie upon some *boundary curve* L of $J + B$. Then let x denote that component of B which contains P, and choose along the curve L a point Q ($\neq P$) sufficiently close to P as to lie also in x. With P and Q as end-points we take an arc PQ such that $C \cdot PQ = P + Q$.[11] The continuum $J + B + PQ$ may now be thought of as the sum of the two continua $J + B$ and $L + PQ$ which have the common *boundary curve* L; hence, by Theorem E_2 of T. I.,[12] it follows that $J + B + PQ$ does not contain a *primitive skew curve.*

Let us next be assured that J is a *boundary curve* of $J + B + PQ$. In view of the manner in which the arc PQ was *added*, it is clear that the point-set $x + PQ$ constitutes a *component* of $B + PQ$, and on comparing it with the component x of B, we observe that these two point-sets both have the same set of limit points on J; that is $F(x + PQ) = F(x)$.[13] Furthermore, the remaining components of $B + PQ$ are each precisely identical with some component of B. Thus it can be seen, by Definition I of T. I., that the *boundary curve* J of $J + B$ is also a *boundary curve* of the continuum $J + B + PQ$.

It is next observed that $C + PQ$ is the sum of the two continua $A + J$ and $J + B + PQ$ which have J as a common *boundary curve.* Then, on employing Theorem E_2 again, we find that $C + PQ$ does not contain a *primitive skew curve*; hence (as this continuum is also *cyclic*) it has a topological image $C' + P'Q'$, $= t(C + PQ)$, in a spherical surface S. The arc $P'Q'$ evidently lies, except for its end-points P' and Q', within some domain D (in S) complementary to C'. But the boundary of D is a simple closed curve Z' which contains P' and is moreover a *boundary curve* of C'. Consequently the point-set $Z = t^{-1}(Z')$ contains the point P and is likewise a *boundary curve* of C, in contradiction to our hypothesis. We conclude then that P cannot lie upon any boundary curve of $J + B$.

LEMMA 3. *Let a and b be distinct points which do not lie together on any*

[11] In taking the arc PQ we may assume that C has been topologically imbedded in a spherical surface S lying in euclidean 3-space and that $PQ - (P + Q)$ is chosen in a domain complementary to S.

[12] Page 828.

[13] For $N \subset M$, $F(N) = \overline{N} \cdot \overline{(M - N)}$.

boundary curve of C; then C contains a subset homeomorphic with one of the following figures:[14]

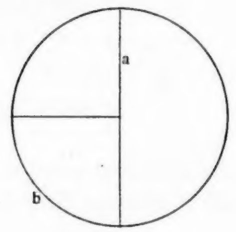


FIG. 3

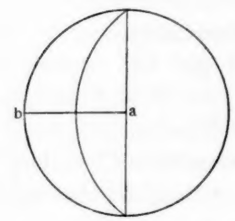


FIG. 4

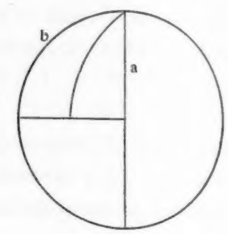


FIG. 5

DEMONSTRATION: As C is a cyclic continuum, it is well-known that the points a and b lie together upon some simple closed curve J in C.

For an arbitrary $\epsilon > 0$, let C be covered by a *finite* collection $\mathfrak{S}^\epsilon$ of its regions[15] U_x of diameter $< \epsilon$. Within each of the regions U_x we choose a point p_x so that

$$(\alpha) \qquad p_x \subset J \quad \text{whenever} \quad U_x \cdot J \neq 0,$$

and

$$(\beta) \qquad p_x \neq p_y \quad \text{when} \quad x \neq y.$$

For each pair of overlapping regions U_x and U_y $(x \neq y)$ one sees by (β) that the points p_x and p_y are distinct, and hence will bound an arc L_{xy} lying in $U_x + U_y$. We set $H = J + \sum_{x,y} L_{xy}$ and proceed to show that each of the arcs L_{xy} in H may be replaced by a peanian subcontinuum N_{xy} of H in a fashion that:

$$\text{(a)} \qquad N_{xy} \supset p_x + p_y,$$

$$\text{(b)} \qquad N_{xy} \subset U_x + U_y,$$

and

$$\text{(c)} \qquad M, = J + \sum_{x,y} N_{xy}, \qquad \text{is a } \textit{linear graph}.$$

First let $M_0 = J$, and then arrange the arcs $\{L_{xy}\}$ in a sequence $\{L_{x_\nu y_\nu}\}$, $(\nu = 1, 2, \cdots, l)$. Next, if M_{i-1} $(1 \leqq i \leqq l)$ is a linear graph lying in H, let $\eta^{(i)} > 0$ be so chosen that if p and q are points of M_{i-1} then:

$$(2) \qquad \rho(p, q) < \eta^{(i)} \quad \text{implies} \quad \rho_{M_{i-1}}(p, q) < \eta,$$[16]

where

$$(3) \qquad \eta = \min \rho[L_{x_\nu y_\nu}, F(U_{x_\nu} + U_{y_\nu})], \qquad \nu = 1, 2, \cdots, l.$$

The point-set $L_{x_i y_j} - M_{i-1}$ consists of an at most countable set of components, the *closures* of which are simple continuous arcs. Those of these arcs which

[14] See Kuratowski's Theorem B, loc. cit., p. 280.

[15] As C is *locally connected*, we may (and do) assume all our *regions* to be *connected*.

[16] The expression $\rho_M(p, q) < \epsilon$ indicates that the points p and q lie together upon an arc of M of diameter $< \epsilon$.

contain either of the points p_{x_i} and p_{y_i}, or which are of diameter $\geqq \eta^{(i)}$, form a *finite* sequence $[A_{i,r}]$, $r = 1, 2, \cdots, r_i$. The remaining arcs constitute a sequence $\{B_{i,s}\}$, $s = 1, 2, \cdots$, which is either *finite* or else:

$$\lim_{s=\infty} \delta(B_{i,s}) = 0. \tag{4}$$

As both of the end-points of $B_{i,s}$ lie at a distance $<\eta^{(i)}$ upon M_{i-1}, we can, according to (2), replace each $B_{i,s}$ by a co-extremal arc $B'_{i,s}$ on M_{i-1} of diameter

$$\delta(B'_{i,s}) < \eta, \qquad s = 1, 2, \cdots. \tag{5}$$

Furthermore, the sequence $\{B'_{i,s}\}$, if not *finite*, may as a consequence of (4) be so chosen that

$$\lim_{s=\infty} \delta(B'_{i,s}) = 0. \tag{6}$$

Then on putting

$$N_{x_i y_i} = \Big(L_{x_i y_i} - \sum_s B_{i,s}\Big) + \sum_s B'_{i,s} \tag{7}$$

it follows from (6) that $N_{x_i y_i}$ is a *peanian* continuum, which moreover *contains the points* p_{x_i} *and* p_{y_i} in satisfaction of requirement (a).

The relations (5) and (7) insure now that every point of $N_{x_i y_i}$ is at a distance $<\eta$ from $L_{x_i y_i}$, so that by (3) we have $N_{x_i y_i} \subset U_{x_i} + U_{y_i}$ (requirement (b)).

Let us next set

$$M_i = M_{i-1} + \sum_{r=1}^{r_i} A_{i,r}, \tag{8}$$

and observe that M_i is a linear graph for each value of $i = 1, 2, \cdots, l$. On considering then the structure of the two continua M_i and $N_{x_i y_i}$ as represented in (7) and (8), one sees that $M_i = M_{i-1} + N_{x_i y_i}$, $i = 1, 2, \cdots, l$; and consequently that

$$M_l = M_0 + \sum_{i=1}^{l} N_{x_i y_i}. \tag{9}$$

In (9) we replace M_0 by J and suppress the subscripts e and i to obtain finally the relation $M = J + \sum_{x,y} N_{xy}$, in conformity with (c).[17]

Let us now make the construction just described for each ϵ_n of a sequence $\{\epsilon_n\}$ that converges to zero; then the relations (α), (β), (a), (b), (c) may henceforth be thought of as holding under the tacit assumption that the U's, p's, N's and M have been replaced respectively by U^n's, p^n's, N^n's, and M^n.

Since $\{\epsilon_n\}$ converges to zero, one may for an arbitrary $\xi > 0$ choose an

[17] The author's construction of the graph M duplicates almost in its entirety a procedure of Mazurkiewicz, *Über nicht plattbare kurven*, Fund. Math., Vol. 20, pp. 282–283.

integer $n(\xi)$ such that if p and q are points of J, and n is taken $>n(\xi)$, then

(10) $$\rho(p, q) < \epsilon_n \quad \text{implies} \quad \rho_J(p, q) < \xi,$$

and

(11) $$3\epsilon_n < \xi.$$

At this point we introduce three propositions, as follows:

PROPOSITION I. *Given $n > n(\xi)$, and P at any point of the point-set $M^n \cdot U^n_\sigma$; then there exists in M^n a peanian continuum X of diameter $<\xi$ that contains both P and p^n_σ.*

Suppose first that $J \supset P$. Then, according to the mode of construction of the graph M^n we conclude from (α) that the point p^n_σ is also on J. Consequently, since $\rho(P, p^n_\sigma) < \epsilon_n$, it follows from (10) that there exists upon J an arc X of diameter $<\xi$ which contains $P + p^n_\sigma$. Furthermore $J \subset M^n$ implies $X \subset M^n$, which establishes the proposition.

Suppose next that $J \not\supset P$. In view, then, of (c), it is clear that $P \subset N^n_{x'y'}$ for some pair of integers (x', y'). Should it now happen that either $x' = \sigma$ or $y' = \sigma$, then by (a′) we see that $N^n_{x'y'}$ also contains the point p^n_σ. Our proposition now follows, on setting $X = N^n_{x'y'}$, since (b) and (11) insure $\delta(X) \leqq \delta(U^n_{x'} + U^n_{y'}) < 2\epsilon_n < \xi$. If, on the other hand, $x' \neq \sigma \neq y'$, it is to be observed that at least one of the regions $U^n_{x'}$ and $U^n_{y'}$, say $U^n_{x'}$, *overlaps* the region U^n_σ. Consequently, the summation $\sum_{x,y} N^n_{xy}$, (which appears in the right member of (c)) must include a continuum $N^n_{\sigma x'}$ associated with (and containing) the points p^n_σ and $p^n_{x'}$.[18] But the point $p^n_{x'}$ is also a point of the continuum $N^n_{x'y'}$; hence $X, = N^n_{\sigma x'} + N^n_{x'y'}$, is a peanian continuum containing $P + p^n_\sigma$, which is of diameter $<\xi$ because of (11) and $\delta(N^n_{\sigma x'} + N^n_{x'y'}) \leqq \delta(U^n_\sigma + U^n_{x'} + U^n_{y'}) \leqq 3\epsilon_n$.

PROPOSITION II. *Given $n > n(\xi)$, and Δ any arc in C having its end-points on the graph M^n; then M^n contains an arc Π, co-extremal with Δ, which is at an upper distance $u(\Pi, \Delta) < \xi$ from Δ.*[19]

Let P_1 and P_2 be the end-points of the arc Δ. Then the covering collection $\mathfrak{S}^{\epsilon_n}$ evidently contains a set of distinct regions $U^n_{\sigma_1}, U^n_{\sigma_2}, \cdots, U^n_{\sigma_r}$, such that

(12) $$U^n_{\sigma_1} \supset P_1, \qquad U^n_{\sigma_r} \supset P_2,$$

(13) $$U^n_{\sigma_i} \cdot U^n_{\sigma_{i+1}} \neq 0 \quad \text{for} \quad 1 \leqq i \leqq r - 1 \quad (\text{when } r > 1),$$

(14) $$U^n_{\sigma_i} \cdot \Delta \neq 0, \qquad 1 \leqq i \leqq r.$$

As a consequence of our hypothesis and relation (12), it follows that $P_1(P_2)$ is a point of $M^n \cdot U^n_{\sigma_1}$ (of $M^n \cdot U^n_{\sigma_r}$); therefore according to Proposition I there

---

[18] We call the reader's attention to the fact that no importance has been attached to the *order* of the subscripts x and y; e.g. the continuum which we denote with $N^n_{\sigma x'}$ might well appear in the summation as $N^n_{x'\sigma}$.

[19] The *upper distance*, $u(M, N)$, of a point-set M from a point-set N is the *upper bound* of the numbers $\rho(m, N)$, where m is a point of M.

exists in M^n a peanian continuum $X_1(X_2)$ of diameter $<\xi$ containing the two points P_1 and $p^n_{\sigma_1}$ (P_2 and $p^n_{\sigma_r}$).

Suppose that $r = 1$. Then the two peanian continua X_1 and X_2 have in common the point $p^n_{\sigma_1} = p^n_{\sigma_r}$, and their sum contains an arc Π having P_1 and P_2 as its end-points. This arc Π is furthermore at an *upper distance* $<\xi$ from Δ; for, an arbitrary point p on Π must lie in X_j for either $j = 1$ or $j = 2$, then

$$\rho(p, \Delta) \leqq \rho(p, P_j) \leqq \delta(X_j) < \xi, \tag{15}$$

which, in view of the *compactness* of Π, implies $u(\Pi, \Delta) < \xi$.

Suppose that $r > 1$. Then, for each value of i $(= 1, 2, \cdots, r-1)$ the relation (13) insures that there is a continuum $N^n_{\sigma_i \sigma_{i+1}}$ appearing in the summation $\sum_{x,y} N^n_{xy}$. If we now put

$$R = X_1 + \sum_{i=1}^{r-1} N^n_{\sigma_i \sigma_{i+1}} + X_2, \tag{16}$$

it follows from $X_1 \supset p^n_{\sigma_1}$, $X_2 \supset p^n_{\sigma_r}$, and (a), that R is *connected* and hence a peanian continuum. Take Π to be any arc lying in R and having P_1 and P_2 as its end-points, and consider an *arbitrary* point p lying on Π. If p should lie in either X_1 or X_2 then, as in (15), we observe that $\rho(p, \Delta) < \xi$. If, on the other hand, $p \not\subset X_j$ $(j = 1, 2)$, then from (16) we conclude that $p \subset N^n_{\sigma_k \sigma_{k+1}}$ for some value of k, $= 1, 2, \cdots, r-1$. Consequently (b) implies either $p \subset U^n_{\sigma_k}$ or $p \subset U^n_{\sigma_{k+1}}$; and in either case $\rho(p, \Delta) < \epsilon_n$ follows directly from (14) and $\delta(U^n_x) < \epsilon_n$. Then (11) infers $\rho(p, \Delta) < \xi$, and on account of the compactness of Π we finally get $u(\Pi, \Delta) < \xi$.

PROPOSITION III. *If, for* $i = 1, 2, \cdots, n, \cdots$, J_{μ_i} *is a boundary curve of the graph* M^{μ_i} *and the sequence* $\{J_{\mu_i}\}$ *converges to a sequential limit* L *of diameter* > 0, *then* L *is a boundary curve of* C.

We first show that the collection $\{J_{\mu_i}\}$ *is equicontinuous.* Suppose on the contrary that this is not so, then from $\{J_{\mu_i}\}$ we can readily choose a subsequence $[J_{n_i}]$, $i = 1, 2, \cdots$, each curve J_{n_i} of which contains a point-pair (P_i, Q_i) such that

1°) the point sequences $[P_i]$ and $[Q_i]$ converge each to a single point T in C; and
2°) the two arcs determined on J_{n_i} by (P_i, Q_i) can be designated by A_i and B_i in a fashion that the arc sequences $[A_i]$ and $[B_i]$ converge respectively to *sequential* limits A and B of diameter >0.

Since C is *cyclic*, there exists in C an arbitrarily small region U containing the point T such that its complement $C - U$ is a *peanian* continuum;[20] and, as both of the continua A and B are of diameter >0, it is clear that U may be taken sufficiently small that

$$(C - U)\cdot A \neq 0 \neq (C - U)\cdot B. \tag{17}$$

[20] Cf. Kuratowski, *Une caractérisation topologique de la surface de la sphère*, Fund. Math., Vol. 13 (1929), pp. 307–318; see *property* (γ) pp. 314–315, and note (2), p. 315.

Next, let W be a subregion of U containing T, for which

$$3\xi = \rho(W, C - U) > 0. \tag{18}$$

An integer ν can now be chosen sufficiently large that we have simultaneously

(i) $n_\nu > n(\xi)$,

(ii) $P_\nu + Q_\nu \subset W$, (a consequence of (1°) and $W \supset T$)

(iii) $(C - U)\cdot A_\nu \neq 0 \neq (C - U)\cdot B_\nu$, (a consequence of (2°) and (17)).

We see by (iii) that the peanian continuum $C - U$ contains an arc Δ having one end-point on A_ν and the other on B_ν. As these end-points both lie upon the graph M^{n_ν}, it follows from Proposition II that M^{n_ν} contains an arc Π, co-extremal with Δ, for which

$$u(\Pi, \Delta) < \xi. \tag{19}$$

Then, since $C - U \supset \Delta$, one concludes from (19) that

$$u(\Pi, C - U) < \xi. \tag{20}$$

Let us now choose upon Π a subarc $\alpha\beta$ which has its end-points α and β lying respectively in A_ν and B_ν, without having any other point on J_{n_ν}. Then (20) implies

$$u(\alpha\beta, C - U) < \xi. \tag{21}$$

P_ν and Q_ν are, according to (ii), points of W and hence are the end-points of an arc Δ' lying in this region. Thus again, by Proposition II, it is to be observed that the graph M^{n_ν} contains an arc Π', co-extremal with Δ', for which

$$u(\Pi', \Delta') < \xi. \tag{22}$$

Then as a consequence of $W \supset \Delta'$ and (22), we obtain

$$u(\Pi', W) < \xi. \tag{23}$$

There exists upon Π' an arc $\alpha'\beta'$ having one end-point α' on the arc $\alpha P_\nu \beta$ (of J_{n_ν}), the other end-point β' on the arc $\alpha Q_\nu \beta$ (of J_{n_ν}), and not touching J_{n_ν} at any other point. The inequality (23) implies then that

$$u(\alpha'\beta', W) < \xi, \tag{24}$$

and it is now clear that the two arcs $\alpha\beta$ and $\alpha'\beta'$ do not have a common point; for, in view of (18), (21), and (24), we see that $\rho(\alpha\beta, \alpha'\beta') > \xi$. Observe finally, however, that the point pair $\alpha + \beta$ separates the point pair $\alpha' + \beta'$ on J_{n_ν}; so that (in the light of lemma 1 of T. I.) J_{n_ν} is *not* a *boundary curve* of M^{n_ν}. This contradiction of our hypothesis establishes the *equicontinuity* of the collection $\{J_{\mu_i}\}$.

In consideration next of the fact that the limit L of the *equicontinuous* collection $\{J_{\mu_i}\}$ *is of diameter* >0, we consult lemma 6 of T. I. in ascertaining that L *is actually a simple closed curve.*

It now remains but to show that L is a *boundary curve* of C. Suppose that this is not the case; then according to lemma 1 of T. I. there exist either

(a°) arcs a_1a_3 and a_2a_4 in C, having no point in common and only their end-points on L, such that $(a_1 + a_3)$ separates $(a_2 + a_4)$ on L;

or if not, then

(b°) arc triods T_1 and T_2 in C such that $T_1 \cdot T_2 = T_1 \cdot L = T_2 \cdot L = a + b + c$, where a, b, c, are the *feet* of the triods T_1 and T_2.

Case 1. *Assume that the situation* (a°) *subsists.* Accordingly, we choose a number $\epsilon > 0$ satisfying the condition

(25) $$5\epsilon < \min. [\rho(a_1a_3, a_2a_4), \rho(a_1, a_3), \rho(a_2, a_4)],$$

and take four regions U_1, U_2, U_3, U_4 in C such that

(26) $$U_i \supset a_i, \qquad \delta(U_i) < \epsilon, \qquad i = 1, 2, 3, 4,$$

and

(27) if b_i is a point of $L \cdot \bar{U}_i$ $(i = 1, 2, 3, 4)$ then the point pair $b_1 + b_3$ *separates* $b_2 + b_4$ on L.

Within each region U_i we choose another region V_i containing the point a_i, such that

(28) $$\rho_i = \rho(V_i, C - U_i) > 0, \qquad i = 1, 2, 3, 4.$$

Then we put

(29) $$3d_i = \rho[a_ia_{i+2} - (V_i + V_{i+2}), L], \qquad i = 1, 2$$

and take

(30) $$\xi = \min (\epsilon, \rho_1, \rho_2, \rho_3, \rho_4, d_1, d_2).$$

An integer h may now be chosen sufficiently large that

(i°) $\mu_h > n(\xi)$,

(ii°) $V_i \cdot J_{\mu_h} \neq 0$, $\qquad i = 1, 2, 3, 4,$

(iii°) $u(J_{\mu_h}, L) < \xi$, and

(iv°) if c_i is a point of $J_{\mu_h} \cdot \bar{U}_i$ $(i = 1, 2, 3, 4)$, then the point-pair $c_1 + c_3$ separates $c_2 + c_4$ on J_{μ_h}.[21]

The relation (ii°) insures that the connected point-set $V_i + a_ia_{i+2} + V_{i+2}$ $(i = 1, 2)$ contains an arc Δ_i having one end-point in $V_i \cdot J_{\mu_h}$ and the other in $V_{i+2} \cdot J_{\mu_h}$; and, in view of (i°), we can, according to Proposition II, approximate the arc Δ_i by a co-extremal subarc Π_i of the graph M^{μ_h} such that

(31) $$u(\Pi_i, \Delta_i) < \xi, \qquad i = 1, 2.$$

---

[21] That the integer h may be taken sufficiently large as to imply (iv°) is assured by the *equicontinuity* of the collection $\{J_{u_i}\}$ together with the relation (27).

There clearly exists upon Π_i $(i = 1, 2)$ a subarc $c_i c_{i+2}$ having only the end-point c_i in $\bar{U}_i \cdot J_{\mu_h}$, and only the end-point c_{i+2} in $\bar{U}_{i+2} \cdot J_{\mu_h}$. Moreover *the arc* $c_i c_{i+2}$ $(i = 1, 2)$ *has no other points than its end-points on the curve* J_{μ_h}. For, suppose, on the contrary, that some *interior* point p of $c_\sigma c_{\sigma+2}$ should lie on J_{μ_h}. Then (iii°) implies

$$\rho(p, L) < \xi, \tag{32}$$

and since $p \subset \Pi_\sigma$, the relation (31) insures

$$\rho(p, \Delta_\sigma) < \xi. \tag{33}$$

According to (33) however, there exists a point $q \subset \Delta_\sigma$ such that

$$\rho(p, q) < \xi. \tag{34}$$

As p does not lie in either U_σ or $U_{\sigma+2}$, we conclude, in view of (28), (30), and (34), that the point q cannot lie in either V_σ or $V_{\sigma+2}$. Hence, since $\Delta_\sigma \subset V_\sigma + a_\sigma a_{\sigma+2} + V_{\sigma+2}$, we see that q is a point of $a_\sigma a_{\sigma+2} - (V_\sigma + V_{\sigma+2})$; which in consideration of (34), implies

$$\rho[p, a_\sigma a_{\sigma+2} - (V_\sigma + V_{\sigma+2})] < \xi. \tag{35}$$

Finally, on combining (30), (32), and (35), one gets

$$\rho[a_\sigma a_{\sigma+2} - (V_\sigma + V_{\sigma+2}), L] < 2\xi \leqq 2d_\sigma; \tag{36}$$

and since the inequality (36) is in contradiction to (29), we must conclude that the arc $c_\sigma c_{\sigma+2}$ has only its end-points c_σ and $c_{\sigma+2}$ on J_{μ_h}.

Let us furthermore be convinced that the two arcs c_1c_3 and c_2c_4 do not overlap one another. From $c_i c_{i+2} \subset \Pi_i$ $(i = 1, 2)$, and (31), we get

$$u(c_i c_{i+2}, \Delta_i) < \xi, \qquad i = 1, 2; \tag{37}$$

which of course implies

$$u[c_i c_{i+2}, (V_i + a_i a_{i+2} + V_{i+2})] < \xi, \qquad i = 1, 2. \tag{38}$$

But, since $U_i \supset V_i$ $(i = 1, 2, 3, 4)$, it follows by (25), (26), and (30) that

$$\rho[(V_1 + a_1a_3 + V_3), (V_2 + a_2a_4 + V_4)] > 3\epsilon \geqq 3\xi. \tag{39}$$

Hence, on combining (38) and (39), we obtain $\rho(c_1c_3, c_2c_4) > \xi$; which insures that the arcs c_1c_3 and c_2c_4 have no points in common.

If one now observes, by (iv°) that the points-pair $c_1 + c_3$ separates the point-pair $c_2 + c_4$ on J_{μ_h}, then it follows that we have finally encountered the contradiction that J_{μ_h} *is not a boundary curve of* M^{μ_k}.

Case 2. *Suppose that situation* (b°) *subsists.* As in Case 1, we are here led in an analogous fashion to the contradiction that a sufficiently large integer k exists for which J_{μ_k} will not be a *boundary curve* of the graph M^{μ_k}. The details of the discussion are left to the reader.

In the light of the contradictions that were encountered in the treatment of

Cases 1 and 2, we conclude that L is necessarily a *boundary curve* of C; thereby completing the demonstration of Proposition III.

We now proceed directly with the proof of the lemma under consideration. If we suppose that the two points a and b lie together on some *boundary curve* J_n of the graph M^n for *every* value of $n = 1, 2, \cdots$; then it is clear that we can choose from the sequence $\{J_n\}$ a subsequence $\{J_{n_i}\}$ converging to a *sequential* limit L of diameter >0. Then, from $J_{n_i} \supset a + b$ $(i = 1, 2, \cdots)$ and $\lim_{i=\infty} J_{n_i} = L$, it immediately follows that $L \supset a + b$. But, according to Proposition III, L is a *boundary curve* of C; thus we have contradicted our hypothesis that a and b do not lie together on any *boundary curve* of C.

There exists then some integer g such that the points a and b *do not lie together upon any boundary curve of the graph M^g*. However, according to Kuratowski,[22] M^g has a topological image $t(M^g)$ in the plane S. If we now recall that the points a and b lie on the simple closed curve J of M^g, then it is apparent that the points $t(a)$ and $t(b)$ lie upon the curve $t(J)$ in $t(M^g)$. The points $t(a)$ and $t(b)$ are furthermore not *conjugate*[23] points of $t(M^g)$. For if such were the case, they would lie together on the frontier F of some maximal domain in S complementary to *that cyclic element of $t(M^g)$ containing* $t(a) + t(b)$. Then on recognizing F as a *boundary curve* of $t(M^g)$, one would conclude (in contradiction of our hypothesis) that $t^{-1}(F)$ is a *boundary curve* of M^g containing $a + b$.

It appears then that $t(a)$ and $t(b)$ are *non-conjugate* points of $t(M^g)$ which lie upon the simple closed curve $t(J)$; hence, by Kuratowski's Theorem B,[24] these two points must lie upon a curve Γ in $t(M^g)$ that is homeomorphic with one of the figures 3, 4, and 5 (where to be sure $t(a)$ and $t(b)$ are situated upon Γ in a manner precisely comparable with the location of the points a and b in the corresponding figure). We now observe in conclusion that the homeomorphic image $t^{-1}(\Gamma)$ of Γ, which lies in M^g and contains $a + b$, is also a subset of C. Q. E. D.

LEMMA 4. *Let P, $\subset C$, be a point that does not lie on any boundary curve of C, and let P' be any point of $C - P$; then there exists a normal representation $C = A + J + B$ such that*

(i) $$A \supset P', \qquad B \supset P, \qquad \textit{and}$$

(ii) *no point of J lies together with P' on any boundary curve of $A + J$.*

DEMONSTRATION: As a consequence of our hypothesis, the points P and P' do not lie together on any *boundary curve* of C. Hence, by Lemma 3, there exists in C a curve Γ homeomorphic with some one of the figures 3, 4, 5; where it is to be understood that the points a and b occurring in the figures are here denoted by P and P'. Now take $\epsilon > 0$, such that if c' is any closed one-cell of Γ that does not contain P, then

(40) $$\rho(c', P) > \epsilon.$$

---

[22] See result (A) in the Introduction.

[23] Cf. Kuratowski, loc. cit. p. 278.

[24] Cf. Kuratowski, loc. cit., p. 280.

According to Lemma 1, there exists a normal representation $C = A + J + B$, such that

$$B \supset P, \tag{41}$$

and

$$\delta(J + B) < \epsilon. \tag{42}$$

From the structure of Γ it is apparent that a *one-cell* c' of Γ which contains the point P' does not also contain P; hence we see from (40) that P and P' are at a distance $> \epsilon$, which implies, in view of the inequality (42), that $A \supset P'$. And this together with (41) insures that the normal representation $A + J + B$ satisfies the condition (i) of our lemma.

We next show that this normal representation also satisfies the condition (ii). Suppose, on the contrary that there is a point Q' on J that lies together with P' on some *boundary curve* J' of $A + J$. Then let H = *that component of* $A \cdot \Gamma$ *which contains* P', and consider the curve $H + J$. It may be observed without difficulty that we can choose within $H + J$ a curve Γ' homeomorphic with one of the figures 3, 4, 5, (where, of course, the points a and b which appear in the figure are now denoted by P' and Q'). We next let $P'Q'$[25] denote any arc having only its end-points P' and Q' on C. It appears then that the point-set $\Gamma' + P'Q'$ is a *primitive skew curve* which, moreover, is contained in

$$A + J + P'Q'.$$

This, however, is altogether impossible. For, it is readily seen that $A + J + P'Q'$ may be represented as the sum of the two peanian continua $(A + J)$ and $(J' + P'Q')$, *each of which has* J' *as a boundary curve*. Furthermore, neither $(A + J)$ nor $(J' + P'Q')$ contains a *primitive skew curve*; hence, by Theorem E_2 of T. I., it follows that $A + J + P'Q'$ does not contain a *primitive skew curve*. We conclude then that the *normal representation* $A + J + B$ satisfies the condition (ii) of our lemma.

Demonstration of the Theorem

Let K be a peanian continuum which is not homeomorphic with a subset of a spherical surface; and let us assume, moreover, that K does not contain a primitive skew curve. Then, according to proposition (B)[26] it immediately follows that K has a cut-point P which is not a *boundary point* of the closure of *every* component of $K - P$. We now proceed to show that this situation implies (in proof of our Theorem) that K must contain a topological image of one of the figures 1 and 2.

Let M be the closure of some component of $K - P$ such that P is *not* a *boundary point* of M. Since $M - P$ is *connected*, we conclude, in view of the

---

[25] See footnote 11.
[26] Cf. Introduction.

well-known structure of a peanian continuum relative to its cyclic elements, that either

Case (I): P lies in some *proper* cyclic element C_1 of M; or

Case (II): P is an *end-point* of M.

THE TREATMENT OF CASE (I): We begin by introducing the following

CONSTRUCTIVE METHOD: Given (1), that the point P lies within, but not upon any *boundary curve* of, a continuum C_n, and (2) a point P_n of $C_n - P$; then we determine in C_n a curve $\Gamma_n + Q_nP_{n+1}$, as follows:

There exists, by Lemma 4, a normal representation $C_n = A_n + J_n + B_n$ such that

(1) $$A_n \supset P_n, \qquad B_n \supset P,$$

and

(2) no point of J_n lies together with P_n on any *boundary curve* of $A_n + J_n$.

We now set $C'_n = J_n + B_n$; then since P does not lie upon any *boundary curve* of C_n it follows, by Lemma 2, that this point does not lie upon any *boundary curve* of C'_n. Hence, according to Lemma 1, there exists for

(3) $$\epsilon_n = \frac{1}{n}\rho(P, J_n)$$

a normal representation $C'_n = A'_n + J'_n + B'_n$, such that

(4) $$B'_n \supset P,$$

and

(5) $$\delta(J'_n + B'_n) < \epsilon_n .$$

From the relations (3), (4), and (5) we get $J_n \subset A'_n$, so that the peanian continuum $A'_n + J'_n$ evidently contains an arc Q_nP_{n+1}, with end-points Q_n and P_{n+1}, such that $J_n \cdot (Q_nP_{n+1}) = Q_n$ and $J'_n \cdot (Q_nP_{n+1}) = P_{n+1}$. As Q_n is a point of the curve J_n, it follows from (2) that P_n and Q_n do not lie together upon any *boundary curve* of the *cyclic* continuum $A_n + J_n$. Therefore, by Lemma 3, there exists in $A_n + J_n$ a topological image Γ_n of one of the figures 3, 4, and 5; where P_n and Q_n occupy positions upon Γ_n comparable with those of the points a and b in the corresponding figure. We now take the sum $\Gamma_n + Q_nP_{n+1}$ as our required curve; for which we evidently have

$$\Gamma_n \cdot (Q_nP_{n+1}) = Q_n .$$

To continue we put

(6) $$C_{n+1} = J'_n + B'_n ;$$

then (4) implies $P \subset C_{n+1}$. Furthermore, *the point P does not lie on any boundary curve of C_{n+1}*; for if so, then by applying Lemma 2, we would encounter the contradiction that P also lies on some *boundary curve* of C_n. Hence, since $P_{n+1} \subset C_{n+1} - P$, we are in position to proceed with the construction of the next curve $\Gamma_{n+1} + Q_{n+1}P_{n+2}$.

Let us now consider the above mentioned cyclic element C_1 of M. As C_1 is a *proper* cyclic element, it follows that $C_1 \neq P$; which allows us to choose P_1 an arbitrary point in $C_1 - P$. Furthermore, since the point P is not a *boundary point* of M, it follows that P cannot lie upon any *boundary curve* of C_1. Hence, we employ our Constructive Method for $n = 1$ and obtain thereby an initial curve $\Gamma_1 + Q_1P_2$; with the observation that in so doing we have also obtained C_2 and P_2, thus setting the stage for the determination of $\Gamma_2 + Q_2P_3$. By continuing this process for the successive values of $n = 2, 3, 4, \cdots$, we generate an infinite sequence of curves $\{\Gamma_n + Q_nP_{n+1}\}$, and then consider their sum $Z = \sum_{n=1}^{\infty} (\Gamma_n + Q_nP_{n+1})$. From the manner in which Z is constructed, it is readily seen that if $u > v$, then

$$(\Gamma_u + Q_uP_{u+1})\cdot(\Gamma_v + Q_vP_{v+1}) = \begin{cases} P_{v+1}, & \text{when } u = v + 1 \\ 0 & \text{when } u > v + 1. \end{cases}$$

Furthermore, on substituting $n - 1$ for n in (5) and (6), we get $\delta(C_n) < \epsilon_{n-1}$, $n = 2, 3, \cdots$; which of course implies $\rho(x_n, P) < \epsilon_{n-1}$ for any point x_n of $\Gamma_n + Q_nP_{n+1}$. Thus, since $\lim_{n=\infty} \epsilon_n = 0$, it follows that P is a limit point of Z, which moreover is the only one that does not belong to Z.

Now let PQ be any arc lying within $\overline{K - M}$ which emanates from the *cut-point* P of M, and then set

$$\Phi = Z + PQ.$$

It is apparent that Φ is a closed point-set; and we proceed to show that Φ contains a subset homeomorphic with one of the figures 1 and 2. Since the curve Γ_n is, for each value of n, homeomorphic with some one of the three figures 3, 4, 5, it follows that there exists an infinite sequence of integers $n_1 < n_2 < \cdots < n_i < \cdots$, such that either

(α) Γ_{n_i} is homeomorphic with figure 3, ($i = 1, 2, \cdots$), or

(β) Γ_{n_i} is homeomorphic with figure 4, ($i = 1, 2, \cdots$), or

(γ) Γ_{n_i} is homeomorphic with figure 5, ($i = 1, 2, \cdots$).

Let us now set

$$G_i = Q_{n_i}P_{n_i+1} + \sum_{v=n_i+1}^{n_{i+1}-1} (\Gamma_v + Q_vP_{v+1}), \qquad i = 1, 2, \cdots;$$

and observe that the two end-points Q_{n_i} and $P_{n_{i+1}}$ of each continuum G_i are joined by an arc $Q_{n_i}P_{n_{i+1}}$ lying in G_i.

Suppose that condition (α) *subsists.* Then on putting

$$\Psi = \sum_{i=1}^{\infty} (\Gamma_{n_i} + P_{n_i}P_{n_{i+1}}) \tag{7}$$

it is a simple matter to identify $\Psi + PQ$ as being a curve that is homeomorphic with figure 1.

Suppose that condition (β) *subsists.* The curve Ψ is here also represented by the summation (7); but, in view of condition (β), $\Psi + PQ$ is in this instance homeomorphic with figure 2.

Suppose that condition (γ) *subsists.* Then let t_i ($i = 1, 2, \cdots$) represent *that unique vertex on the curve* Γ_{n_i} *which is of order* 4; and let $Q_{n_i}t_i$ denote that arc upon Γ_{n_i} which has Q_{n_i} and t_i as its end-points *and which has no vertex of* Γ_{n_i} *as an interior point.* We then set $\Psi' = \sum_{i=1}^{\infty} [(\Gamma_{n_i} - Q_{n_i}t_i) + Q_{n_i}P_{n_{i+1}}]$, and observe finally that $\Psi' + PQ$ is homeomorphic with figure 1. Having now completed the treatment of Case (I), let us pass to

THE TREATMENT OF CASE (II): We choose in M an arbitrary arc B which emanates from the point P; and let $\{C_n\}$ be the sequence consisting of *all* of those cyclic elements of M such that for each value of n

(i) $B \cdot C_n$ consists of *more* than a single point.

and

(ii) $B \cdot C_n$ is *not* a subset of any boundary curve of M.

The sequence $\{C_n\}$ *must contain an infinite subsequence converging to* P; for if not, there would exist a subarc β of B, containing P, such that if C were *any* cyclic element of M having *more* than one point on β, then $\beta \cdot C$ would be a subset of some *boundary curve* of C. Consequently, according to Definition II of T. I., β would be a *boundary arc* of M and we would thereby encounter the contradictory inference that P is a *boundary point* of M.

In view of (i) it is a simple matter to choose, for each value of n, a subarc P_nQ_n of B having its end-points P_n and Q_n on C_n, such that $B \cdot C_n \subset P_nQ_n$, and P_n and Q_n lie upon B in the *order* P_nQ_nP. Suppose now that for some number $\epsilon > 0$, it is true that the inequality $\rho(C_n, P) < \epsilon$ implies that the points P_n and Q_n *lie together upon a boundary curve* of C_n. Then, if $u_1, u_2, \cdots, u_i, \cdots$, be an integer sequence consisting of *all* values of n for which the cyclic element C_n is at a distance $< \epsilon$ from P, it is clear that each subarc $P_{u_i}Q_{u_i}$ of B can now be replaced by a *co-extremal* arc $P_{u_i}t_iQ_{u_i}$ *which lies along some boundary curve of* C_{u_i}. Accordingly we set

$$B' = (B - \textstyle\sum_{i=1}^{\infty} P_{u_i}Q_{u_i}) + \sum_{i=1}^{\infty} P_{u_i}t_iQ_{u_i},$$

and observe that *B' is also an arc.* Furthermore, if B'' is any subarc of B', which contains P and is of diameter $< \epsilon$, *then B'' is a boundary arc of M.* But this contradicts our hypothesis that *P is not a boundary point of M*; consequently, we conclude that there exist cyclic elements C_n lying *arbitrarily* close to P such that the points P_n and Q_n *do not lie upon any boundary curve of* C_n.

In view of the preceding considerations, it is now a simple matter to choose from $\{C_n\}$ a subsequence $\{C_{n_i}\}$, $i = 1, 2, \cdots$, such that

(i) $B \cdot C_{n_i} \subset P_{n_i}Q_{n_i}$, $\quad i = 1, 2, \cdots$;

(ii) P_{n_i} and Q_{n_i} do not lie upon any boundary curve of C_{n_i}, $\quad i = 1, 2, \cdots$,

(iii) $\lim_{i=\infty} C_{n_i} = P$,

(iv) the *point order* $P_{n_1}Q_{n_1}P_{n_2}Q_{n_2} \cdots P_{n_i}Q_{n_i} \cdots P$ is preserved upon the arc B.

As a consequence of (i), our Lemma 3 applies here for $C = C_{n_i}$, $a = P_{n_i}$, $b = Q_{n_i}$; and we are assured of the existence in C_{n_i} ($i = 1, 2, \cdots$) of a curve Γ_{n_i} homeomorphic with one of the figures, 3, 4, 5. Let $Q_{n_i}P_{n_{i+1}}$ denote that unique subarc of B that joins the two distinct points Q_{n_i} and $P_{n_{i+1}}$; then from (i) and (iv) we are assured that $\Gamma_{n_i} \cdot (Q_{n_i}P_{n_{i+1}}) = Q_{n_i}$, $\Gamma_{n_{i+1}} \cdot (Q_{n_i}P_{n_{i+1}}) = P_{n_{i+1}}$, $i = 1, 2, \cdots$. On now putting $Z = \sum_{i=1}^{\infty} (\Gamma_{n_i} + Q_{n_i}P_{n_{i+1}})$ one sees, as a consequence of (ii) and (iv), that P is the sole limit point of Z not belonging to Z. Accordingly, we choose PQ an arbitrary arc in $\overline{K - M}$ emanating from the point P, and set

$$\Phi = Z + PQ.$$

From this point on one may proceed precisely as in Case 1 to show that Φ (and hence K) contains subset homeomorphic with either figure 1 or figure 2. This concludes the proof of our THEOREM.

In view finally of the obvious fact that neither of the figures 1 and 2 has a topological image in a spherical surface, we can state as follows:

FUNDAMENTAL THEOREM. *A necessary and sufficient condition that a peanian continuum be homeomorphic with a subset of the surface of a sphere is that it contains neither a primitive skew curve nor a topological image of either of the figures 1 and 2.*

THE UNIVERSITY OF MICHIGAN
ANN ARBOR, MICHIGAN

Annals of Mathematics
Vol. 38, No. 3, July, 1937

ALEXANDERSCHER UND GORDONSCHER RING UND IHRE ISOMORPHIE

von Hans Freudenthal

(Received November 10, 1936)

J. W. Alexander [3] hat eine Produktbildung für die Elemente der von ihm selbst [1, 2, 3] und von A. Kolmogoroff [1, 2, 3] entdeckten *oberen* Bettischen Gruppen[1] konstruiert; je ein Element der oberen Bettischen Gruppe ρ-ter und σ-ter Dimension liefern als Produkt ein Element der oberen Bettischen Gruppe $(\rho + \sigma)$-ter Dimension. I. Gordon [1] hat im Außenraum eines in der d-dimensionalen Sphäre liegenden Polytops P eine Multiplikation der *unteren* d.h. gewöhnlichen) Bettischen Gruppen definiert, bei der je einem Element der Bettischen Gruppe ρ-ter and σ-ter Dimension des Außenraumes ein Element der Bettischen Gruppe $(\rho + \sigma - d)$-ter Dimension des Außenraumes als Produkt zugeordnet wird. Nun weiß man aber durch J. W. Alexander [1, 2] und A. Kolmogoroff [1, 4, 5], daß die ρ-te obere Bettische Gruppe von P isomorph der $(d - \rho)$-ten unteren Bettischen Gruppe des Außenraumes von P ist. Das Verhalten der Dimensionszahlen läßt vermuten, daß sich diese Isomorphie auch auf die Produktibildungen erstreckt, daß also die Alexandersche Multiplikation in P und die Gordonsche im Außenraum von P im Wesentlichen miteinander identisch sind. Diese Vermutung werden wir bestätigen (allgemein für beliebige abgeschlossene Teilmengen P der Sphäre und beliebige Koeffizientenbereiche). Unser Isomorphiesatz verhält sich zu den Dualitätssätzen Jordanscher Richtung etwa wie der letzte Teil der Čechschen Untersuchungen [1], der die Isomorphie zwischen Schnittbildung und Produktbildung in Mannigfaltigkeiten feststellt, zu den Dualitätssätzen Poincaréscher Richtung. Unsere Methode ähnelt der Čechschen.

Der größte Teil unserer Arbeit[1a] dient der Einführung einer zweckmäßigen Symbolik.

1. Ein *Simplexgitter* ist ein System von „Ecken", in dem für gewisse endliche Eckenmengen ein alternierendes „Produkt" $[a_0 a_1 \cdots a_\rho]$, ρ-dimensionales *Simplex* genannt, derart erklärt ist, daß jede echte Teilmenge der Ecken eines Simplexes wieder ein Simplex erzeugt. Die Eckenmenge wird im Folgenden stets als endlich vorausgesetzt.

Unterer (oberer) ρ-dimensionaler Komplex $k_\rho (k^\rho)$ heißt vorläufig jede Linear-

[1] Bei J. W. Alexander „duale Bettische Gruppen" genannt. Wir folgen A. Kolmogoroff in unsern Bezeichnungen.

[1a] Eine vorläufige Mitteilung der Ergebnisse der vorliegenden Arbeit findet sich in der Note: *Über Mannigfaltigkeiten und ihre Abbildungen*, Proceed. Akad. Amsterdam *40* (1937), 54–60.

form in den ρ-dimensionalen Simplexen mit Koeffizienten aus der kompakten (abzählbaren diskreten) additiven abelschen Gruppe $\Gamma(\Delta)$; dabei seien Γ und Δ Charakteregruppen [Pontrjagin 1] mod 1[2] voneinander. Die $k_\rho(k^\rho)$ bilden eine kompakte (abzählbare diskrete) Gruppe $K_\rho(K^\rho)$. Das Simplexgitter wird, wo es nötig ist, in Klammern angedeutet. Bei den Simplexen wird die Dimension, je nachdem in was für einem Komplex sie gerade auftreten, ebenfalls durch einen unteren oder oberen Index angedeutet.

$t^\rho t_\rho$ soll dann und nur dann $+1(-1)$ sein, wenn beide „Faktoren" dasselbe (entgegengesetzte) Simplex darstellen, sonst 0. Bei Verwendung von Kommutativitäts- und Distributivitätsvorschriften ist dann das Produkt $k^\rho k_\rho$ allgemein als reelle Zahl mod 1 erklärt. Auf die Art werden K^ρ und K_ρ dual zueinander (d.h. Charakteregruppen voneinander).

2. R und S seien zwei Simplexgitter und $\mathfrak{f}$ ein stetiger Homomorphismus

$$\mathfrak{f}K_\rho(R) \subset K_\sigma(S);$$

die Gleichung

$$k^\sigma(S)(\mathfrak{f}k_\rho(R)) = (k^\sigma(S)\mathfrak{f})k_\rho(R)$$

(identisch in k^σ and k_ρ) definiert dann einen Homomorphismus

$$k^\sigma(S)\mathfrak{f} \subset k^\rho(R)$$

(also in umgekehrter Richtung; es ist zweckmäßig, hier das Funktionszeichen *hinter* das Argument zu schreiben). Unsere Bemerkung läßt sich natürlich umkehren.

Hin und wieder werden wir einen Komplex durch die ihm zugehörige (alternierende) Koeffizientenfunktion φ andeuten;[3] das Zeichen $\mathfrak{f}$ tritt dann auch vor oder hinter die Koeffizientenfunktion.

Beispiele für Homomorphismen $\mathfrak{f}$:

a) $\rho = \sigma$.— $\mathfrak{s}$ sei eine *simpliziale Abbildung* von R in S. $\mathfrak{s}$ induziert einen stetigen Homomorphismus von $K_\rho(R)$ in $K_\rho(S)$, der auch $\mathfrak{s}$ genannt wird, $\mathfrak{s}[a_0 \cdots a_\rho] = [\mathfrak{s}a_0 \cdots \mathfrak{s}a_\rho]$, also auch einen Homomorphismus von $K^\rho(S)$ in $K^\rho(R)$; dabei ist $\varphi\mathfrak{s}[a_0 \cdots a_\rho] = \varphi[\mathfrak{s}a_0 \cdots \mathfrak{s}a_\rho]$, also jedes Simplex geht in die Summe seiner Urbilder über.

b) $\rho = \sigma$.— S sei eine *Zerschlagung* von R, d. h. es gebe eine Zuordnung $\mathfrak{z}$ (mit noch besonderen Eigenschaften), die jedem Simplex t_ρ von R eine Summe Σu_ρ verschiedener Simplexe u von S zuordnet (wobei jedes absolute u_ρ in genau einer Summe auftritt). $\mathfrak{z}$ induziert einen stetigen Homomorphismus $\mathfrak{z}K_\rho(R) \subset K_\rho(S)$, also auch einen Homomorphismus $K^\rho(S)\mathfrak{z} \subset K^\rho(R)$. Dabei ist $\varphi\mathfrak{z}(u^\rho) = \psi(t^\rho) = \Sigma\varphi(u^\rho)$, welche Summe zu erstrecken ist über die u^ρ von $t^\rho = \Sigma u^\rho$.

---

[2] Als Bereich der Produkte kann man auch allerlei andere Gruppen wählen [H. Freudenthal 2, Nr. 16, 47].

[3] Nur hin und wieder. Meistens führt die Verwendung von Funktionen statt Komplexen zu großen Komplikationen.

c) $\rho = \sigma + 1$, $R = S$.— $\mathfrak{r}$ ordne jedem k_ρ seinen „*unteren*" *Rand* $\mathfrak{r}k_\rho$ zu, also

$$\mathfrak{r}[a_0 \cdots a_\rho] = \Sigma_i(-1)^i[a_0 \cdots \hat{a}_i \cdots a_\rho].$$

(Wegzulassende Ecken sind mit einem Zirkumflex versehen.) $k^\rho\mathfrak{r}$ ist dann der *obere Rand* [Kolmogoroff 1, 2] oder die „Ableitung" [Alexander 1, 2, 3] von k^ρ. Man hat

$$\varphi\mathfrak{r}[a_0 \cdots a_{\rho+1}] = \Sigma_i(-1)^i\varphi[a_0 \cdots \hat{a}_i \cdots a_{\rho+1}].$$

Man bemerkt, daß $\mathfrak{s}$ und $\mathfrak{r}$ miteinander *vertauschbar* sind; zu den nicht explicit genannten Eigenschaften der Zerschlagung gehört auch die Vertauschbarkeit von $\mathfrak{r}$ und $\mathfrak{z}$.

3. Ein *unterer* (*oberer*) *Zyklus* $z_\rho(z^\rho)$ ist ein $k_\rho(k^\rho)$ mit $\mathfrak{r}k_\rho = 0$ $(k^\rho\mathfrak{r} = 0)$; ein *unterer* (*oberer*) *Rand* $h_\rho(h^\rho)$ ist ein Element von $\mathfrak{r}K_\rho(K^\rho\mathfrak{r})$. (Bei Alexander heißen die z^ρ „exakt" und die h^ρ „abgeleitet"). Die $z_\rho(z^\rho)$ bilden eine abgeschlossene Untergruppe $Z_\rho(Z^\rho)$ von $K_\rho(K^\rho)$ und die $h_\rho(h^\rho)$ eine abgeschlossene Untergruppe $H_\rho(H^\rho)$ von $Z_\rho(Z^\rho)$. Die *untere* (*obere*) *Bettische* $B_\rho(B^\rho)$ ist die Faktorgruppe $Z_\rho/H_\rho(Z^\rho/H^\rho)$. Aus Pontrjaginschen Sätzen [1] folgt, daß B_ρ und B^ρ dual zueinander sind [Alexander 1, Kolmogoroff 1, 3].

Das Simplexgitter, auf das sich die Gruppen beziehen, deuten wir, wo es nötig ist, in Klammern an.

Wegen der letzten Bemerkung von 3 gilt: $\mathfrak{s}$ induziert einen stetigen Homomorphismus $\mathfrak{s}B_\rho(R) \subset B_\rho(S)$ und einen Homomorphismus von $B^\rho(S)\mathfrak{s} \subset B^\rho(R)$. Entsprechendes gilt für $\mathfrak{z}$. (Daß bei den oberen Bettischen Gruppen der Homomorphismus gerade in umgekehrter Richtung induziert wird, macht sie in manchen Fällen weniger handlich als die unteren.)

4. R sei ein *kompakter metrischer Raum*, ϑ_n eine monotone positive Nullfolge, R_n eine endliche Teilmenge von R, von der jeder Punkt von R einen Abstand $<\vartheta_n$ habe, weiter sei $R_n \subset R_{n+1}$. R_n wird ein Simplexgitter durch die Festsetzung: $[a_0 \cdots a_\rho]$ ist für jede Teilmenge $a_0, \cdots, a_\rho$ von R_n definiert, deren Durchmesser $< \epsilon_n$ ist; dabei sei ϵ_n eine positive monotone Nullfolge mit $\epsilon_n \geqq \epsilon_{n+1} + 2\vartheta_n$. Zu jedem n gibt es eine simpliziale Abbildung $\mathfrak{s}_n R_{n+1} \subset R_n$, die jeden Punkt um höchstens ϑ_n verrückt. Nach der letzten Bemerkung von 4 bilden also die Bettischen Gruppen $B_\rho(R_n)$ bzw. $B^\rho(R_n)$ eine G_n-adische bzw. G_n-ale Folge [H. Freudenthal 1; 2, Nr. 2, 11] mit den Abbildungen $\mathfrak{s}_n B_\rho(R_{n+1}) \subset R_n$ bzw. $B^\rho(R_n)\mathfrak{s}_n \subset B(R_{n+1})$. Der G_n-adische bzw. G_n-ale Limes der Folge heiße *untere bzw. obere Bettische Gruppe* $B_\rho(R)$ *bzw.* $B^\rho(R)$ des kompakten Raumes R. Aus allgemeinen Sätzen [Freudenthal 2, Nr. 8, 13, 39] folgt die Unabhängigkeit dieser Gruppen von der speziellen Wahl der ϑ_n usw. Aus allgemeinen Dualitätssätzen über Gruppenfolgen [Freudenthal 2, Nr. 20, 21] folgt die Dualität von $B_\rho(R)$ und $B^\rho(R)$.

5. Eine stetige Abbildung $\mathfrak{f}$ eines kompakten Raumes R in einen kompakten Raum S induziert einen stetigen Homomorphismus der G_n-adischen Folge

$B_\rho(R_n)$ in die G_n-adische $B_\rho(S_n)$ [Freudenthal 2, Nr. 19] und damit auch der Limesgruppe $B_\rho(R)$ in die Limesgruppe $B_\rho(S)$; ebenso einen stetigen Homomorphismus der G_n-alen Folge $B^\rho(S_n)$ in die G_n-ale Folge $B^\rho(R_n)$ und damit auch der Limesgruppe $B^\rho(S)$ in die Limesgruppe $B^\rho(R)$.

6. Sei jetzt R ein *im Kleinen kompakter Raum*, also Vereinigung einer aufsteigenden Folge kompakter Räume R_n. Wendet man 5 an auf die eineindeutige Abbildung von R_n in R_{n+1} (für alle n), so kann man die Bettischen Gruppen $B^\rho(R)$ bzw. $B_\rho(R)$ als G_n-adischen bzw. G_n-alen Limes der $B^\rho(R_n)$ bzw. $B_\rho(R_n)$ definieren.

7. Die Bettischen Gruppen eines *endlichen Polytops* P darf man in der Weise berechnen, daß man die Polytopecken zu Ecken eines Simplexgitters R macht und die Eckenmengen, die konkrete Simplexe erzeugen, Simplexe von R erzeugen läßt; die Bettischen Gruppen des Gitters R stimmen mit denen von P überein.

Aus Dualitätsgründen genügt es, diese Tatsache für die unteren *oder* die oberen Bettischen Gruppen zu beweisen; für die unteren ist sie wohlbekannt [z. B. Freudenthal 2, Nr. 41]. Trotzdem wollen wir einen selbständigen Beweis für die oberen Bettischen Gruppen skizzieren:

Die durch ein abstraktes Simplex in P definierte Punktmenge „konkretes Simplex" soll durch ein vorgesetztes $\mathfrak{m}$ angedeutet werden.

Wir metrisieren P so, daß die und nur die Punktmengen von R, die Simplexe erzeugen, einen Durchmesser <1 haben. Von der endlichen Menge S, $R \subset S \subset P$ habe jeder Punkt von P einen Abstand $<\epsilon$; Simplexe u^ρ von S seien dadurch charakterisiert, daß sie einen Durchmesser $<\epsilon$ haben. Man darf annehmen, daß eine Zerschlagung $\mathfrak{z}t_\rho = \Sigma u_\rho$ existiert, bei der die in $\mathfrak{z}t_\rho$ auftretenden $\mathfrak{m}u_\rho$ das zugehörige $\mathfrak{m}t_\rho$ doppelpunktfrei überdecken. Ferner darf man die Existenz einer simplizialen Abbildung $\mathfrak{s}S \subset R$ voraussetzen, bei der jedes $\mathfrak{m}t$ auf sich selbst abgebildet wird; $\mathfrak{s}$ wird als Produkt $\mathfrak{p}^{(1)}\mathfrak{p}^{(2)} \cdots \mathfrak{p}^{(d)}$ genommen, in dem $\mathfrak{p}^{(\rho)}$ innerhalb jedes $\mathfrak{m}t_\rho$ eine Projektion von einem geeigneten Punkt im Innern auf den Rand darstellt. Daß $\mathfrak{s}\mathfrak{z}$ in $B^\rho(R)$ den identischen Automorphismus erzeugt, ist (bei geeigneter Wahl der Projektionszentren) klar. Wir haben nur noch zu zeigen, daß auch $\mathfrak{z}\mathfrak{s}$ in $B^\rho(S)$ den identischen Automorphismus induziert, also daß in S aus $\varphi^\rho\mathfrak{z}\mathfrak{s} \sim 0$ folgt: $\varphi^\rho \sim 0$. Für jedes t^ρ setzen wir $\varphi_{t^\rho}(u^\rho) = \varphi(u^\rho)$, wenn $\mathfrak{m}u^\rho \subset \mathfrak{m}t^\rho$, sonst $= 0$. φ_{t^ρ} hat die Koeffizientensumme 0, berandet in $\mathfrak{m}t^\rho$ also ein $\chi_{t^\rho}^{\rho-1}$; $\varphi_{(1)}^\rho = \varphi^\rho - \Sigma_{t^\rho}\chi_{t^\rho}^{\rho-1}\mathfrak{r}$ verschwindet für alle u^ρ mit $\mathfrak{m}u^\rho$ in geeignetem $\mathfrak{m}t^\rho$. Wir setzen weiter $\varphi_{(1)t^{\rho+1}}(u^\rho) = \varphi_{(1)}(u^\rho)$ für $\mathfrak{m}u^\rho \subset \mathfrak{m}t^\rho$, sonst $= 0\cdot$ $\varphi_{(1)t^{\rho+1}}$ ist in $\mathfrak{m}t^{\rho+1}$ ein oberer ϵ-Zyklus, berandet dort also ein $\chi_{(1)t^{\rho+1}}(u^{\rho-1})$; $\varphi_i(u^\rho) = \varphi_i(u^\rho) - \Sigma\chi_{(1)t^{\rho+1}}^{\rho-1}\mathfrak{r}$ verschwindet für alle u^ρ mit $\mathfrak{m}u^\rho$ in geeignetem $\mathfrak{m}t^{\rho+1}$. So fährt man durch alle Dimensionen fort.

8. R sei ein Simplexgitter (Simplexe t), S seine *baryzentrische* Zerschlagung, d.h. Ecken von S sind die Punktmengen $a_0, \cdots, a_\rho$ (auch $a_{0\ldots\rho}$ genannt), für die $[a_0 \cdots a_\rho]$ erklärt ist und nicht verschwindet, Simplexe u_ρ von S sind die

Ausdrücke, die sich nach evtl. Umordnung schreiben lassen als $[a_{\mathfrak{n}_0} a_{\mathfrak{n}_1} \cdots a_{\mathfrak{n}_\rho}]$, so daß die Indexmengen $\mathfrak{n}_0, \mathfrak{n}_1, \cdots, \mathfrak{n}_\rho$ eine aufsteigende Folge bilden.

$\mathfrak{z}[a_0 a_1 \cdots a_\rho] = \Sigma \mp [a_{\nu_0} a_{\nu_0 \nu_1} \cdots a_{\nu_0 \nu_1 \cdots \nu_\rho}]$; die Summe ist über alle Permutationen der $0, 1, \cdots, \rho$ zu erstrecken, das Vorzeichen $\mp$ ist das der betreffenden Permutation.

R sei eine d-dimensionale Pseudomannigfaltigkeit; alle z_d sind von der Form ${}_a\Sigma \mp t_d$; das Vorzeichen, mit dem ein t_d in dieser Summe auftritt, heiße $\pi(t_d)$. Wir betrachten die Abbildung

$$\mathfrak{d}[a_0 \cdots a_\rho] = \Sigma \pi(a_0 \cdots a_d)[a_{0 \cdots \rho} a_{0 \cdots \rho, \rho+1} \cdots a_{0 \cdots \rho, \rho+1, \cdots d}];$$

hier steht also rechts die zu $[a_0 \cdots a_\rho]$ gehörige Dualzelle. Wir beweisen nun[4]

$$\mathfrak{r}(\mathfrak{d}k) = (-1)^{\rho+1} \mathfrak{d}(k\mathfrak{r}). \tag{1}$$

Wir brauchen den Beweis bloß für $k = t = [a_0 \cdots a_\rho]$ zu führen. In

$$\mathfrak{r}\Sigma \mp [a_{0 \cdots \rho} \cdots a_{0 \cdots \rho \cdots i} \cdots a_{0 \cdots \rho \cdots i \cdots d}]$$

tritt jedes $[a_{0 \cdots \rho} \cdots \hat{a}_{0 \cdots \rho \cdots i} \cdots a_{0 \cdots \rho \cdots i \cdots d}]$ für $i \leqq d$ zweimal und zwar mit entgegengesetztem Zeichen auf; das folgt für $i < d$ durch Vertauschung von i und $i + 1$, für $i = d$ aus der Tatsache, daß R eine Pseudomannigfaltigkeit ist. Wir haben also einerseits

$$\mathfrak{r}\mathfrak{d}[a_0 \cdots a_\rho] = \Sigma \mp [a_{0 \cdots \rho, \rho+1} \cdots a_{0 \cdots \rho, \rho+1 \cdots d}],$$

andererseits

$$\begin{aligned} \mathfrak{d}([a_0 \cdots a_\rho]\mathfrak{r}) &= (-1)^{\rho+1} \mathfrak{d}\Sigma [a_0 \cdots a_\rho a_{\rho+1}] \\ &= (-1)^{\rho+1} \Sigma \mp [a_{0 \cdots \rho, \rho+1} \cdots a_{0 \cdots \rho, \rho+1 \cdots d}], \end{aligned}$$

woraus die zu beweisende Formel folgt.

$\mathfrak{d}$ bildet also Z^ρ homomorph in Z_ρ (mit dem Koeffizientenbereich Δ) ab, ebenso B^ρ in B_ρ.

9. Von nun an seien die a_i von R in einer *festen Reihenfolge* angeordnet);[5] ein Simplex heißt *normal*, wenn seine Ecken in der gegebenen Reihenfolge aufgeschrieben sind.

---

[4] Wir behandeln k also sowohl als oberen als auch als unteren Komplex. D.h. wir verlangen von den Koeffizientengruppen von jetzt an nicht mehr, daß sie dual, sondern daß sie identisch sind. Ganz exakt wäre es, hier einen Operator einzuführen, der einen Koeffizientenbereich mit seinem dualen vertauscht, und einen Operator, der einen unteren durch den zugehörigen oberen Komplex ersetzt (und umgekehrt). Wir verzichten darauf.

[5] Die Verwendung einer festen Reihenfolge rührt von E. Čech und H. Whitney her. Daß die wesentlichen Begriffsbildungen von dieser Willkür unabhängig sind, ergibt sich bei uns als Nebenresultat. Was eigentlich hinter dieser festen Reihenfolge steckt, zeigt die simpliziale Abbildung $\mathfrak{o}$ aus Nr. 9 der vorliegenden Arbeit; unter Verwendung von $\mathfrak{o}$ könnte man übrigens auch den Unabhängigkeitsbeweis direkt führen. Zusatz bei der Korrektur: Den Gedanken, daß hinter der festen Anordnung der Ecken die simpliziale Abbildung $\mathfrak{o}$ steckt, hat (unabhängig von der vorliegenden Arbeit) auch Herr S. Lefschetz gehabt; siehe seinen Vortrag *The Role of Algebra in Topology*, Bull. Am. Math. Soc. (1937), pp. 345–359, den ich inzwischen im Manuskript kennen lernte.—Bemerkt sei noch, daß man mittels $\mathfrak{o}$ das Alexandersche auch auf das Cartesische Produkt zurückführen kann.

$\mathfrak{o}$ sei die simpliziale Abbildung, die jedes $a_{\mathfrak{n}_0\mathfrak{n}_1\cdots\mathfrak{n}_\rho}$ ($\mathfrak{n}_0 \subset \mathfrak{n}_1 \subset \cdots \subset \mathfrak{n}_\rho$) in das a_σ überführt, das in der gegebenen Reihenfolge das höchste ist unter allen a_τ mit $\tau \epsilon \mathfrak{n}_\rho$. In nicht verschwindende t_ρ gehen die und nur die u_ρ über, die sich schreiben lassen als $[a_{0\cdots\lambda} a_{0\cdots\lambda,\lambda+1} \cdots a_{0\cdots\lambda,\lambda+1,\cdots\lambda+\rho}]$, wobei (in der gegebenen Reihenfolge) a_ν höher steht als a_μ, sobald $\lambda \leqq \mu < \nu$.

$$\text{(2)} \qquad \mathfrak{o}\mathfrak{d}[a_0 \cdots a_\rho] = \Sigma \mp [a_\rho \cdots a_d],$$

wo nur über normale Simplexe summiert wird und das Vorzeichen $\pi(a_0 \cdots a_d)$ ist.

10. Für normale Simplexe $t^\rho = [a_0 \cdots a_\rho]$, $t^\sigma = [b_\rho \cdots b_{\rho+\sigma}]$ wird $t^\rho \cdot t^\sigma$ definiert als $[a_0 \cdots a_\rho b_{\rho+1} \cdots b_{\rho+\sigma}]$, falls $a_\rho = b_\rho$ ist, sonst $= 0$ (also auch $= 0$, wenn $a_0, \cdots a_\rho, b_{\rho+1}, \cdots, b_{\rho+\sigma}$ kein Simplex erzeugen); das ist die Alexandersche *Produktbildung* [3]. In naheliegender Weise ist dann $k^\rho \cdot l^\sigma$ erklärt, falls die Koeffizienten von k^ρ bzw. l^σ Koeffizientenbereichen entnommen sind, zwischen denen eine Multiplikation besteht; k^ρ und l^σ sind dabei als lineare Kombination *normaler* Simplexe anzusetzen.

Von Alexander übernehmen wir, daß damit eine Multiplikation zwischen den Elementen von B^ρ und B^σ definiert ist (die Produkte sind Elemente von $B^{\rho+\sigma}$).

Bei simplizialen Abbildungen, die die Reihenfolge nicht zerstören, geht das Produkt zweier Bildsimplexe (Komplexe, Elemente der Bettischen Gruppen) eines Simplexgitters in das Produkt der Urbildsimplexe (-komplexe usw.) des Urbildgitters über (für Zerschlagungen gilt das Entsprechende nicht). Die Betrachtungen von 4–6 lassen sich also dahin ausdehnen, daß die Produktbildung auch für die B^ρ beliebiger kompakter und im Kleinen kompakter Räume definiert (und bei Berücksichtigung der Topologie der B^ρ) stetig ist. Ebenso zeigen die Betrachtungen von 7, daß die Multiplikation der B^ρ eines Polytops als die im zugehörigen Simplexgitter zu berechnen ist.

11. P sei ein Polytop in der d-dimensionalen Homologiesphäre E; die Teilung von E sei so gewählt, daß P aus konkreten Simplexen, $\mathfrak{m}t$, dieser Teilung bestehe; die Ecken von E seien so angeordnet, daß erst die von P und dann die übrigen kommen; ferner soll P jedes Simplex von E enthalten, von dem es alle Eckpunkte enthält. Die Ecken der baryzentrischen Unterteilung (Simplexe: u), die *nicht* in P liegen, und die aus ihnen erzeugten baryzentrischen Simplexe bilden ein Polytop Q. Von einem t, das nicht ganz in P liegt, sagen wir, es liege in $E \setminus P$. Von einem k, dessen sämtliche t bzw. u in $E \setminus P$ bzw. Q liegen, sagen wir, es liege selbst in $E \setminus P$ bzw. Q. Ist $t \subset E \setminus P$, so ist $\mathfrak{d}t \subset Q$, und umgekehrt. Wir definieren

$$\mathfrak{a} = \mathfrak{r}\mathfrak{d}.$$

z^ρ sei ein oberer Zyklus in P. Dann ist $z^\rho\mathfrak{r} \subset E \setminus P$, also $\mathfrak{d}(z^\rho\mathfrak{r}) \subset Q$, also nach (1) auch $\mathfrak{a}z^\rho = \mathfrak{r}\mathfrak{d}z^\rho \subset Q$.

z^ρ sei obendrein ~ 0 in P. Dann ist $z^\rho = k^{\rho-1}\mathfrak{r} + k^\rho$ mit $k^{\rho-1} \subset P$, $k^\rho \subset E \setminus P$.

Also $\mathfrak{a}z^\rho = \mathfrak{r}\mathfrak{d}(k^{\rho-1}\mathfrak{r}) + \mathfrak{r}\mathfrak{d}k^\rho$. Da wegen (1) der erste Summand verschwindet, ist $\mathfrak{a}z^\rho$ unterer Rand von $\mathfrak{d}k^\rho$, d.h. eines Komplexes aus Q.

Beides zusammen ergibt: $\mathfrak{a}$ erzeugt einen Homomorphismus $\mathfrak{a}B^\rho(P) \subset B_{d-\rho-1}(Q)$, von dem wir zeigen werden, daß er ein "Isomorphismus auf" ist (mit der üblichen Modification im Falle $d - \rho - 1 = 0$):

$z_{d-\rho-1}$ sei ein unterer Zyklus in Q, den wir, da E eine Homologiemannigfaltigkeit ist, in der Form $\mathfrak{d}z^{\rho+1}$, $z^{\rho+1} \subset E \setminus P$ ansetzen dürfen. In E berandet $\mathfrak{d}z^{\rho+1}$ ein $\mathfrak{d}k^\rho$; läßt man aus k^ρ alle Simplexe von $E \setminus P$ weg, so entsteht z^ρ, ein oberer Zyklus in P, und es ist $\mathfrak{a}z^\rho \sim z_{d-\rho-1}$ in Q.

Sei $\mathfrak{a}z^\rho \sim 0$ in Q. Dann ist $\mathfrak{r}\mathfrak{d}z^\rho \sim 0$ in Q, und es gibt ein $k^\rho \subset E \setminus P$ mit $\mathfrak{r}\mathfrak{d}z^\rho = \mathfrak{r}\mathfrak{d}k^\rho$. Wegen (1) ist dann $\mathfrak{d}((z^\rho - k^\rho)\mathfrak{r}) = 0$, also $(z^\rho - k^\rho)\mathfrak{r} = 0$, also $z^\rho - k^\rho = k^{\rho-1}\mathfrak{r}$. Läßt man in dieser Gleichung alle Simplexe von $E \setminus P$ weg, so heißt das: $z^\rho \sim 0$ in P.

$\mathfrak{a}$ vermittelt also einen Isomorphismus zwischen $B^\rho(P)$ und $B_{d-\rho-1}(Q)$ (Alexander-Pontrjaginscher Dualitätssatz).

12. Von zwei Komplexen in einer Homologiemannigfaltigkeit, deren Schnitt man zu bilden hat, setzt man voraus, daß der eine in Dualzellen, der andere in ursprünglichen Simplexen gegeben ist; der Schnittkomplex ergibt sich dann in Simplexen der baryzentrischen Unterteilung. Die Koeffizienten entstammen Bereichen zwischen denen eine Multiplikation erklärt ist.

k_ρ, l_σ seien zwei Komplexe in E, $\rho + \sigma \geqq d$,[6] und sei $k_\rho = \mathfrak{d}k^{d-\rho}$. Der *Schnitt* $\mathfrak{d}k^{d-\rho} \times l_\sigma$ beider drückt sich durch die $u_{\rho+\sigma-d}$ aus. Wir brauchen ihn nur für Dualzellen und Simplexe zu erklären:[7]

$$(\mathfrak{d}[a_0 \cdots a_{d-\rho}]) \times [a_0 \cdots a_{d-\rho} \cdots a_\sigma] = \Sigma \mp [a_{0\cdots d-\rho}\, a_{0\cdots d-\rho,\nu_1} \cdots a_{0\cdots d-\rho,\nu_1\cdots\nu_{\sigma+\rho-d}}], \tag{3}$$

wo über alle Permutationen $\nu_1, \cdots, \nu_{\sigma+\rho-d}$ der $d - \rho + 1, \cdots, \sigma$ zu summieren und für $\mp$ das Vorzeichen der betreffenden Permutation zu nehmen ist,

$$(\mathfrak{d}[a_0 \cdots a_{d-\rho}]) \times [b_0 \cdots b_\sigma] \qquad \text{sonst} = 0.$$

Nimmt man $t^{d-\rho}$ und t_σ normal an, so ist

$$\mathfrak{o}(\mathfrak{d}t^{d-\rho} \times t_\sigma) = [a_{d-\rho}\, a_{d-\rho+1} \cdots a_\sigma] \tag{4}$$

im ersten Fall, sonst 0.

13. Die Gordonsche Operation, $\otimes$ genannt, ist so definiert: je einem Element von $B_{d-\sigma-1}(Q)$ und $B_{d-\rho-1}(Q)$, bzw. repräsentiert durch die Zyklen $z_{d-\sigma-1}$ und $z_{d-\rho-1}$ in Q, ist zugeordnet das Element von $B_{d-\rho-\sigma-1}(Q)$, das repräsentiert wird durch $\mathfrak{r}(k_{d-\sigma} \times k_{d-\rho})$, wo $\mathfrak{r}k_{d-\sigma} = z_{d-\sigma-1}$ und $\mathfrak{r}k_{d-\rho} = z_{d-\rho-1}$ ist. Nach 11

[6] Für $\rho + \sigma < d$ setzt man den Schnitt bekanntlich 0. Im Folgenden lassen wir diesen (trivialen) Fall stets unberücksichtigt.

[7] Siehe J. H. C. Whitehead, Annals of Math. (2) *33* (1932), 521–524.

dürfen wir $z_{d-\sigma-1}$ bzw. $z_{d-\rho-1}$ in der Gestalt $\mathfrak{a}z^\sigma$ bzw. $\mathfrak{a}z^\rho$ annehmen, z^σ, $z^\rho \subset P$. $z_{d-\sigma-1} \otimes z_{d-\rho-1}$ läst sich also berechnen als $\mathfrak{r}(\mathfrak{d}z^\sigma \times \mathfrak{d}z^\rho)$. Da $\mathfrak{o}$ (wegen der Voraussetzung über die Anordnung der Ecken) eine simpliziale Abbildung von Q in sich darstellt, können wir hier auch schreiben

$$(5) \qquad \mathfrak{a}z^\sigma \otimes \mathfrak{a}z^\rho \sim \mathfrak{r}\mathfrak{o}(\mathfrak{d}z^\sigma \times \mathfrak{o}\mathfrak{d}z^\rho).$$

und auf diese Form die Regel (4) enwenden.

Seien $t^\rho = [a_0 \cdots a_\rho]$, $t^\sigma = [a_\rho \cdots a_{\rho+\sigma}]$ normale Simplexe. Dann ist wegen (2)

$$\mathfrak{o}\mathfrak{d}t^\rho = \Sigma\pi(a_0 \cdots a_d)[a_\rho \cdots a_d]$$

(nur über normale Simplexe zu summieren), also wegen (4)

$$\mathfrak{o}(\mathfrak{d}t^\sigma \times \mathfrak{o}\mathfrak{d}t^\rho) = \Sigma\pi(a_0 \cdots a_d)[a_{\rho+\sigma} \cdots a_d].$$

Andererseits ist wegen (2)

$$\mathfrak{o}\mathfrak{d}(t^\rho \cdot t^\sigma) = \Sigma\pi(a_0 \cdots a_d)[a_{\rho+\sigma} \cdots a_d],$$

während für andersartige Simplexe beide Ausdrücke verschwinden. Also

$$\mathfrak{o}(\mathfrak{d}t^\sigma \times \mathfrak{o}\mathfrak{d}t^\rho) = \mathfrak{o}\mathfrak{d}(t^\rho \cdot t^\sigma).$$

Diese Gleichung bleibt gültig, wenn man die t durch Komplexe ersetzt; unter Berücksichtigung von (5) erhält man also

$$\mathfrak{a}z^\sigma \otimes \mathfrak{a}z^\rho \sim \mathfrak{a}(z^\rho \cdot z^\sigma),$$

die zu beweisende Isomorphie zwischen Gordonscher und Alexanderscher Operation.

14. Die Beziehung $\mathfrak{a}$ war von der vorgelegten Reihenfolge der Ecken unabhängig; dasselbe gilt von der Gordonschen Operation. Also liefert auch die Alexandersche Operation bei jeder Anordnung der Ecken dasselbe Resultat für die Produkte von Elementen aus $B^\rho(P)$ und $B^\sigma(P)$.

Sei nun P eine beliebige abgeschlossene Teilmenge von E. Nach 10 hat die Alexandersche Operation für die $B^\rho(P)$ einen Sinn; P ist dabei aufzufassen als R_n-adischer Limes einer absteigenden Folge von Polytopen P_n und $B^\rho(P)$ als G_n-aler Limes der $B^\rho(P_n)$, zwischen denen (für jedes n) die Operation erklärt ist. Die zugehörigen Q_n erzeugen R_n-al (als ihre Vereinigung) Q, die Komplementärmenge von P, und dabei erzeugen die $B_{d-\rho-1}(Q)$ R_n-al $B_{d-\rho-1}(Q)$ (siehe 6); die für jedes n existierende Gordonsche Operation erzeugt dabei eine "Gordonsche Operation" zwischen den $B_{d-\rho-1}(Q)$. Wegen der Isomorphie beider Operationen für jedes n erhält man so die gewünschte Isomorphie beider Operationen hinsichtlich $B^\rho(P)$ bzw. $B_{d-\rho-1}(Q)$, die wir beweisen wollten.

LITERATUR

J. W. Alexander: 1. *On the chains of a complex and their duals.* Proc. Nat. Acad. USA *21* (1935), 509-511.

2. *On the ring of a compact metric space.* Proc. Nat. Acad. USA *21* (1935), 511-512.

3. *On the connectivity ring of an abstract space.* Annals of Math. (2) *37* (1936), 698-708.

E. Čech: 1. *Multiplications on a complex.* Annals of Math. (2) *37* (1936), 681–697.

Hans Freudenthal: 1. *Die R_n-adische Entwicklung von Räumen und Gruppen.* Proc. Akad. Amsterdam *38* (1935), 414–418.

2. *Entwicklungen von Räumen und ihren Gruppen.* Compositio Math. *4* (1936), 145–238.

I. Gordon: 1. *On intersection invariants of a complex and its complementary space.* Annals of Math. (2) *37* (1936), 519–525.

A. Kolmogoroff: 1. *Über die Dualität im Aufbau der kombinatorischen Topologie.* Rec. Math. Moscou *43* (1936), 97–102.

2. *Les groupes de Betti des espaces localement bicompacts.* C. R. Paris *202* (1936), 1144–1147.

3. *Propriétés des groupes de Betti des espaces localement bicompacts.* C. R. Paris *202* (1936), 1325–1327.

4. *Les groupes de Betti des espaces métriques.* C. R. Paris *202* (1936), 1558–1560.

5. *Cycles relatifs.* C. R. Paris *202* (1936), 1641–1643.

L. Pontrjagin: 1. *Topological commutative groups.* Annals of Math. (2) *35* (1934), 361–388.

Amsterdam (Holland).

Annals of Mathematics
Vol. 38, No. 3, July, 1937

UN THÉORÈME SUR L'HOMOTOPIE

Par Samuel Eilenberg

(Received October 8, 1936)

1. Étant donnés deux espaces métriques X et Y, on désigne par Y^X la classe de toutes les transformations continues de X en des sous-ensembles de Y; lorsque Y est borné ou bien X compact, on considère Y^X comme un espace métrique, en posant pour $f_1, f_2 \epsilon Y^X$:

$$| f_1 - f_2 | = \sup_{x \epsilon X} | f_1(x) - f_2(x) |.$$

Deux transformations f_1, $f_2 \epsilon Y^X$ sont dites *homotopes*, quand il existe une transformation $g \epsilon Y^{X \times [0,1]}$ [1] telle que $g(x, 0) = f_1(x)$ et $g(x, 1) = f_2(x)$ pour $x \epsilon X$.

Y est dit *rétracte*[2] de X, lorsque $Y \subset X$ et qu'il existe une fonction $f \epsilon Y^X$ telle que $f(y) = y$ pour $y \epsilon Y$.

Y s'obtient de X par une *déformation continue dans* Z, lorsque $X + Y \subset Z$ et qu'il existe une fonction $g \epsilon Z^{X \times [0,1]}$ telle que $g(x, 0) = x$ pour $x \epsilon X$ et $g(X, 1) = Y$; en d'autres termes, qu'il existe une fonction $f \epsilon Z^X_{\text{où}}$ $f(X) = Y$ homotope à la transformation $f_0 \epsilon Z^X$ de X en X par l'identité (à savoir où $f(x) = g(x, 1)$ et $f_0(x) = g(x, 0)$).

Y est dit *rétracte de* X *par déformation*,[3] lorsque $Y \subset X$ et qu'il existe une fonction $g \epsilon X^{X \times [0,1]}$ telle que $g(x, 0) = x$ pour $x \epsilon X$, $g(X, 1) = Y$ et $g(y, 1) = y$ pour $y \epsilon Y$; en d'autres termes, s'il existe une fonction $f \epsilon X^X$ rétractant X en Y et homotope à la transformation $f_0 \epsilon X^X$ de X en X par l'identité (à savoir où $f(x) = g(x, 1)$ et $f_0(x) = g(x, 0)$).

M. W. Hurewicz[4] a introduit récemment la notion de type d'homotopie. On dit que deux espaces compacts X et Y ont le même *type d'homotopie*,[5] lorsqu'il existe deux transformations $f \epsilon Y^X$ et $g \epsilon X^Y$ telles que les transformations $gf \epsilon X^X$ et $fg \epsilon Y^Y$ sont homotopes respectivement à l'identité de X en X et de Y en Y.

2. Quand on cherche à appliquer ces notions à l'étude de la structure homotopique d'un espace compact, les problèmes suivants se posent:

---

[1] $X \times [0, 1]$ désigne le *produit cartésien* (produit combinatoire) de X et de l'intervalle clos $[0, 1]$.

[2] K. Borsuk, Fund. Math. 17 (1931), pp. 152–170.

[3] K. Borsuk, Fund. Math. 21 (1933), p. 91.

[4] W. Hurewicz, *Beiträge zur Topologie der Deformationen* I, Proceed. Akad. Amsterdam, 38(1935), pp. 112–119; II, ibid., pp. 521-528; III, ibid., vol. 39(1936), pp. 117–125; IV, ibid., pp. 215-224. Nous les désignerons dans les renvois ultérieurs par D. I, D. II, D. III et D. IV.

[5] D. III, p. 124.

Trouver les propriétés d'un espace compact X pour lequel

(α_n) *il existe un sous-ensemble compact Y de X de dimension $\leqq n$ qui est un rétracte par déformation de X;*

(β_n) *il existe un sous-ensemble compact Y de X de dimension $\leqq n$ en lequel X soit déformable par une déformation continue (dans X);*

(γ_n) *il existe un espace compact Y (assujetti au besoin à certaines propriétés locales) de dimension $\leqq n$ et dont le type d'homotopie soit le même que celui de X.*

Le but de cet article est l'étude du cas $n = 1$. Il sera demontré qu'en admettant certaines hypothèses sur X, les propriétés (α_1), (β_1), (γ_1) sont équivalentes et on établira une condition nécessaire et suffisante pourqu'elles aient lieu.

3. Soit G un groupe (discret). Nous appellerons *rétracte* de G[6] tout sous-groupe G' de G pour lequel il existe une transformation homomorphe h de G en G' telle que $h(x) = x$ pour $x \in G'$.

Nous allons considérer une fonction $\tau(G)$[7] qui fait correspondre à tout groupe G un entier non négatif ou ∞, défini comme il suit:

$\tau(G) \geqq n$ équivaut à l'existence d'un sous-groupe G' de G qui est 1° rétracte de G, 2° engendré par n générateurs libres.

Nous aurons recours au théorème suivant

(1) *Si G' est un rétracte de G, les deux groupes étant libres et à n générateurs, on a $G' = G$.*

C'est une conséquence immédiate du théorème suivant:

G étant un groupe libre à un nombre fini de générateurs et h une transformation homomorphe de G qui n'est pas une isomorphie, les groupes G et $h(G)$ ne sont pas isomorphes.[8]

4. Nous avons introduit récemment[9] un coefficient de multicohérence $r(X)$ dont nous ferons intervenir dans la suite les deux propriétés suivantes, se rapportant aux continus X localement connexes pour les dimensions 0 et 1.[10]

(1) *$r(X) \geqq n$ équivaut à l'existence d'un polyèdre (topologique) Y de dimension 1, rétracte de X,*[11] *et tel que $b_1(Y) = n$.*

(2) $$r(X) = \tau[\pi_1(X)],^{12}$$

où $\pi_1(x)$ est le groupe de Poincaré de X.

---

[6] D. IV, p. 220.

[7] S. Eilenberg, *Sur les espaces multicohérents* I, Fund. Math. 27 (1936), p. 775.

[8] W. Magnus, Math. Ann. 111 (1935), p. 276, aussi F. Levi, Math. Zeitschrift 37 (1933), p. 90–97.

[9] S. Eilenberg, l.c., pp. 153–190.

[10] S. Lefschetz, Ann. of Math., 35 (1934), p. 119; C. Kuratowski, Fund. Math. 24 (1935), p. 269. La connection locale pour la dimension 0 n'est autre que la connection locale tout court.

[11] S. Eilenberg, l.c., p. 173.

[12] ibid., p. 175.

L'hypothèse que X est un continu localement connexe pour les dimensions 0 et 1 implique que $\pi_1(X)$ est discret avec un nombre fini de générateurs et de relations définissantes.

Quant au nombre $b_1(Y)$, qui entre dans l'énoncé (1), il suffira de savoir que lorsque Y est un polyèdre de dimension 1 (le seul cas qui nous intéresse ici), ce nombre est égal au premier nombre de Betti de Y.[13] Par conséquent
(3) *Y étant un polyèdre connexe de dimension 1 avec $b_1(Y) = n$, $\pi_1(Y)$ est un groupe à n générateurs libres.*

5. X et Y étant des continus localement connexes pour les dimensions 0 et 1, toute fonction $f \in Y^X$ détermine une transformation homomorphe du groupe $\pi_1(X)$ en sous-groupe de $\pi_1(Y)$. Cette homomorphie reste la même si l'on passe de f à une autre transformation $f_1 \in Y^X$ homotope à f. On montre sans peine que:
(1) *Si $f \in Y^X$ est une rétraction de X en $Y \subset X$, $\pi_1(Y)$ peut être considéré comme un sous-groupe de $\pi_1(X)$ et la transformation homomorphe correspondant à f est une rétraction de $\pi_1(X)$ en $\pi_1(Y)$.*
(2) *Si Y s'obtient de X par une déformation continue (dans X), le groupe $\pi_1(X)$ est isomorphe à un sous-groupe de $\pi_1(Y)$.*

6. Nous désignerons par $\pi_n(X)$ le n-ième groupe d'homotopie de X au sens de M. W. Hurewicz.[14] Le groupe $\pi_1(X)$ est donc considéré comme premier groupe d'homotopie. Pour $n \geqq 2$, le cas $\pi_n(X) = 0$[15] est le seul dont nous aurons à nous occuper. Ce cas est caractérisé par la condition que toute transformation $f \in X^{S_n}$[16] soit *inessentielle*, c.à.d. homotope à une transformation constante.[14]

Les espaces tels que $\pi_n(X) = 0$ pour tout $n \geqq 2$ sont dits *asphériques*.[17] Nous ferons l'usage dans la suite du théorème suivant:
(1) *Soient: X un continu de dimension finie, localement connexe pour les dimensions 0 et 1; Y un continu asphérique et localement connexe pour toutes les dimensions;[10] f_1, $f_2 \in Y^X$ deux transformations déterminant la même transformation homomorphe de $\pi_1(X)$ en un sous-groupe de $\pi_1(Y)$. Alors les transformations f_1 et f_2 sont homotopes.*[18]

Nous montrerons d'abord que
(2) *Le théorème (1) reste vrai lorsqu'on y remplace X par un continu qui est un rétracte absolu de voisinage[19] (et dont la dimension est arbitraire).*

---

[13] K. Borsuk et S. Eilenberg, Fund. Math. 26 (1936), p. 221.
[14] D. I., p. 114.
[15] Nous écrivons $\pi_n(X) = 0$ pour "le groupe $\pi_n(X)$ se compose d'élément neutre."
[16] S_n désigne la surface sphérique à n dimensions.
[17] D. IV, p. 215.
[18] D. IV, p. 219.
[19] Un espace compact est dit *rétracte absolu de voisinage* (K. Borsuk, Fund. Math. 19 (1932), p. 222), lorsque tout espace Y, dans lequel on peut plonger X topologiquement,

Supposons X plongé dans le cube fondamental Q_ω de l'espace de Hilbert. X étant un rétracte absolu de voisinage, il existe un ensemble ouvert $U \supset X$ et une fonction $\varphi \in X^U$ telle que $\varphi(x) = x$ pour $x \in X$.

Il existe aussi un polyèdre (de dimension finie) $P \subset U$ et une fonction continue $\psi \in P^X$ telle que pour tout $x \in X$ le segment rectiligne $\overline{x, \psi(X)}$ soit contenu dans U. Par conséquent, en désignant par f_0 la transformation de X en X par l'identité, les fonctions $f_0 \in U^X$ et $\psi \in U^X$ sont homotopes et il en est de même des fonctions $f_0 \in X^X$ et $\varphi\psi \in X^X$, puisque $\varphi f_0 = f_0$.

Les fonctions $f_1\varphi \in Y^P$ et $f_2\varphi \in Y^P$ déterminent la même transformation homomorphe de $\pi_1(P)$ en un sous-groupe de $\pi_1(Y)$, car les fonctions $f_1 \in Y^X$ et $f_2 \in Y^X$ déterminent la même transformation de $\pi_1(X)$ en un sous-groupe de $\pi_1(Y)$. Le polyèdre P étant de dimension finie, nous pouvons appliquer (1) et en déduire l'homotopie des fonctions $f_1\varphi \in Y^P$ et $f_2\varphi \in Y^P$, donc aussi celle des fonctions $f_1\varphi\psi \in Y^X$ et $f_2\varphi\psi \in Y^X$. La transformation $\varphi\psi \in X^X$ étant homotope à la transformation de X en X par l'identité, il en résulte l'homotopie des fonctions $f_1 \in Y^X$ et $f_2 \in Y^X$, c.q.f.d.

7. Lemme. *Soient: X un continu quelconque, localement connexe pour les dimension 0 et 1, et Y un sous-continu de dimension 1 de X localement connexe (pour la dimension 0). Il existe alors un polyèdre topologique (connexe et de dimension 1) $Y_1 \subset Y$ qui s'obtient de Y par une déformation continue dans X.*

Démonstration. D'après un théorème de M. S. Mazurkiewicz,[20] il existe pour tout $\eta > 0$ un polyèdre topologique $Y_1 \subset Y$ et une fonction continue φ telle que $\varphi(Y) = Y_1$ et $| y - \varphi(y) | < \eta$ pour tout $y \in Y$.

Posons $f_0(y) = y$ pour tout $y \in Y$. Comme $Y \subset X$, on a donc $f_0 \in X^Y$, $f_0\varphi \in X^Y$ et $| f_0 - f_0\varphi | < \eta$.

L'espace X étant localement connexe pour les dimensions 0 et 1, l'espace X^Y est[21] localement connexe (pour la dimension 0). Par conséquent on peut choisir $\eta > 0$ assez petit pour que f_0 et $f_0\varphi$ soient homotopes. Y_1 s'obtient donc de Y par une déformation continue dans X.

8. Théorème. *Les quatres propriétés suivantes sont équivalentes pour tout continu X qui est un rétracte absolu de voisinage:*[19]

(α_1) *il existe un ensemble compact $Y \subset X$ de dimension $\leqq 1$ qui est un rétracte de X par déformation;*

(β_1) *il existe un ensemble compact $Y \subset X$ de dimension $\leqq 1$ qui s'obtient de X par une déformation continue (dans X);*

---

admet un sous-ensemble ouvert $U \supset X$ qui se laisse rétracter en X. Les rétractes absolus de voisinage sont *localement contractiles* (ibid., p. 236–7), donc aussi localement connexes pour toutes les dimensions. Dans le cas où ils sont de dimension finie, ils coïncident avec les espaces localement contractiles (ibid., p. 240) et aussi avec les espaces localement connexes pour toutes les dimensions (C. Kuratowski, l.c.).

[20] S. Mazurkiewicz, Fund. Math. 20 (1933), p. 281.

[21] C. Kuratowski, l.c., p. 285.

(γ_1) *il existe un continu localement connexe (en dimension* 0) *Y de dimension* $\leqq 1$ *dont le type d' homotopie est le même que celui de X;*
(δ_1) $\pi_1(X)$ *est un groupe libre à un nombre fini de générateurs et* $\pi_n(X) = 0$ *pour* $n \geqq 2$.

DÉMONSTRATION. Nous allons établir les implications:

$$(\alpha_1) \to (\beta_1) \to (\delta_1) \to (\alpha_1) \text{ et } (\alpha_1) \to (\gamma_1) \to (\delta_1).$$

(α_1) $\to$ (β_1). Celà est évident, puisqu'un rétracte par déformation s'obtient de X par une déformation continue (dans X).

(β_1) $\to$ (δ_1). On peut admettre en vertu du lemme que Y est un polyèdre topologique de dimension 1.

Le groupe $\pi_1(Y)$ est un groupe libre (à un nombre fini de générateurs) et le groupe $\pi_1(X)$ est isomorphe en vertu de 5(2) à un sous-groupe de $\pi_1(Y)$. Par conséquent $\pi_1(X)$ est un groupe libre, en tant que sous-groupe d'un groupe libre.[22]

Pour montrer que $\pi_n(X) = 0$ pour tout $n \geqq 2$, considérons une transformation $f \in X^{S_n}$ quelconque. La transformation f est donc selon (β_1) homotope à une transformation $f_1 \in X^{S_n}$ telle que $f_1(S_n) \subset Y$. Or, on a $\pi_n(Y) = 0$, puisque $n \geqq 2$ et Y est polyèdre de dimension 1.[23] La transformation f_1, et par conséquent aussi f, est inessentielle.

(δ_1) $\to$ (α_1). Soit n le nombre des générateurs libres du groupe $\pi_1(X)$. On a donc $r[\pi_1(X)] \geqq n$ et, en vertu de 4(2), $r(X) \geqq n$. D'après 4(1), il existe par conséquent un polyèdre (topologique) Y de dimension 1, qui est un rétracte de X et pout lequel $b_1(Y) = n$. Soit $f \in Y^X$ la rétraction de X en Y. Cette fonction détermine selon 5(1) une rétraction de $\pi_1(X)$ en son sous-groupe $\pi_1(Y)$. Or, les deux groupes étant libres et à n générateurs (le premier par hypothèse et le second en vertu de 4(3)), on conclut de 3(1) que $\pi_1(X) = \pi_1(Y)$ et que cette rétraction est une identité. La fonction $f \in X^X$ détermine donc la transformation identique de $\pi_1(X)$ en lui-même. En posant $f_0(x) = x$ pour tout $x \in X$, on trouve que les fonctions $f_0, f \in X^X$ déterminent la même transformation homomorphe de $\pi_1(X)$ (notamment la transformation par l'identité). Il en résulte en vertu de 6(2) que f_0 et f sont homotopes, donc que Y est un rétracte de X par déformation.

(α_1) $\to$ (γ_1). Soit Y un rétracte de X par déformation. Il existe donc une fonction $f \in X^X$ qui effectue la rétraction de X en Y et qui est homotope à la transformation de X en X par l'identité. En considérant les fonctions $f \in Y^X$ et $g \in X^Y$ où $g(y) = y$ pour tout $y \in Y$, on trouve que 1° $gf \in X^X$ coïncide avec $f \in X^X$ et par conséquent est homotope à la transformation de X en X par l'identité, 2° $fg \in Y^Y$ est elle-même une identité, puisque $f(y) = y$ et $g(y) = y$ pour tout $y \in Y$. Les espaces X et Y ont donc le même type d'homotopie.

(γ_1) $\to$ (δ_1). X et Y ayant le même type d'homotopie, les nombres de Betti

---

[22] O. Schreier, Abh. Hamburg Sem. 5 (1927), p. 161.
[23] D. IV, p. 216.

correspondants sont égaux,[24] par conséquent le premier nombre de Betti de Y est fini. Comme Y est un continu localement connexe et de dimension 1, il en résulte que Y est localement contractile. Donc 1° le groupe $\pi_1(Y)$ est libre, 2° $\pi_n(Y) = 0$ pour $n \geqq 2$.[23] Or, $\pi_n(X)$ est isomorphe à $\pi_n(Y)$[24] pour tout n, ce qui achève la démonstration.

9. Remarquons enfin que l'on peut remplacer dans le théorème établi tout à l'heure l'hypothèse que le continu X est un rétracte absolu de voisinage par celle que X est un polyèdre connexe, donc par un cas particulier.[25] En revanche, on n'aura même pas besoin d'admettre dans ce cas que ce polyedre soit fini, de sorte que le théorème pourra être établi aussi pour des polyèdres connexes infinis, même de dimension non bornée. En particulier, on peut prendre comme X un domaine ouvert situé dans un espace euclidien.

Pour démontrer le théorème ainsi modifié, il suffit de disposer des énoncés analogues à 4(1), 4(2) et 6(2). Les deux premiers sont établis dans mon article: *Sur les espaces multicohérents II* (Fund. Math. 29 (1937) p. 114, th. 2 et 4). Quant à la proposition 6(2), elle se démontre exactement comme pour les polyèdres finis.[26]

VARSOVIE.

---

[24] D. III, p. 125.
[25] Cf. renvoi 21.
[26] D. IV, p. 217.

ANNALS OF MATHEMATICS
Vol. 38, No. 3, July, 1937

SUFFICIENT CONDITIONS BY EXPANSION METHODS FOR THE PROBLEM OF BOLZA IN THE CALCULUS OF VARIATIONS[1]

BY WILLIAM T. REID

(Received October 22, 1936)

1. **Introduction.** In 1913 Levi [13][2] established the sufficiency of the well-known conditions (see, for example, [5], p. 128) for a strong relative minimum in the non-parametric plane problem of the calculus of variations *without the use of a field of extremals.* We shall speak of Levi's method as an expansion method since it is an extension of Legendre's idea of expanding the total variation.[3] Described briefly, the method involves certain inequalities on the Weierstrass E-function which are implied by the usual strengthened Weierstrass condition, together with certain elementary integral inequalities. It also utilizes the Clebsch [Jacobi] transformation of the second variation. A very simple proof of this transformation for the problem of Lagrange has been given by Bliss ([1], p. 691; [2], p. 738).

For partial differential equations and more general functional equations one does not have available imbedding theorems corresponding to those known for ordinary differential equations. Consequently, it is to be expected that Levi's method, which avoids the use of such imbedding theorems, will prove fruitful in the study of more complicated problems of the calculus of variations.[4] The writer plans to consider this question in a later paper.

In the present paper Levi's method is applied to the general problem of Bolza in non-parametric form. This problem has been treated rather extensively within recent years by the Weierstrass field method (see the papers of Bliss, Morse, Hestenes, Hu and Reid [19] listed in the bibliography). The extension of the expansion method to the problem of Bolza is not trivial. In particular, it involves the proof of certain inequalities for the Weierstrass E-function. These inequalities are established in Section 4. Section 3 is devoted to an extension of an auxiliary theorem on accessory extremals due to Hestenes. It is to be remarked that the details of Levi's method have been considerably simplified in the present paper. The inequalities of Section 5 are

[1] Presented to the American Mathematical Society, Sept. 10, 1935.

[2] Numerals in square brackets refer to the bibliography at the end of this paper.

[3] In an earlier paper [12] Levi stated a sufficiency theorem for a strong relative minimum and attempted to prove it by expansion methods. The error in his work was pointed out by Hahn [8], who also showed by an example that the theorem as stated by Levi was false. Malnate [14] has extended the results of Levi [13] to the problem in three-space. It is to be remarked that the criticism made by Duren [6] of Levi[13] and Malnate [14] is invalid.

[4] Miranda has recently applied Levi's method to the double integral problem of the calculus of variations [Mem. R. Accad. d'Italia, vol. 5 (1934), pp. 159–172].

due to McShane, who suggested to the author their advantages over certain other integral inequalities previously used in the proof of the sufficiency theorem of this paper. Theorem 6.1 of this paper is the same as Theorem 9.2 of Hestenes [9] and Theorem 21.1 of Bliss [4]. The other sufficiency theorems of Hestenes and Bliss may be derived as consequences of this theorem, since in each case it may be proved that the hypotheses imply the positiveness of the second variation.

In a later paper the author will give sufficient conditions by expansion methods for the problem of Bolza in parametric form.

2. **Formulation of the problem.** The problem of Bolza is that of finding in a class of arcs

$$y_i(x) \qquad (i = 1, \cdots, n;\ x_1 \leqq x \leqq x_2) \tag{2.1}$$

satisfying differential equations and end-conditions

$$\varphi_\beta[x, y, y'] = 0 \qquad (\beta = 1, \cdots, m < n), \tag{2.2}$$

$$\psi_\mu[x_1, y(x_1), x_2, y(x_2)] = 0 \qquad (\mu = 1, \cdots, p < 2n + 2), \tag{2.3}$$

one which minimizes a given functional

$$J = g[x_1, y(x_1), x_2, y(x_2)] + \int_{x_1}^{x_2} f[x, y, y']\, dx. \tag{2.4}$$

In the present paper derivatives of $f[x, y, r]$, $\varphi_\beta[x, y, r]$, $g[x_1, y_{i1}, x_2, y_{i2}]$, $\psi_\mu[x_1, y_{i1}, x_2, y_{i2}]$ with respect to their arguments are indicated by writing these variables as subscripts. The tensor analysis summation convention is used throughout. If $z \equiv (z_i)$, the norm $\sqrt{(z_i z_i)}$ is denoted by $\| z \|$. The particular sets $z_i = 0$, $z_{ij} = 0$ $(i, j = 1, \cdots, n)$ are denoted by (0_i) and (0_{ij}), respectively.

We suppose that we have given an open region $\mathfrak{R}$ of points $[x, y, r] = [x, y_1, \cdots, y_n, r_1, \cdots, r_n]$ in which the functions $f[x, y, r]$, $\varphi_\beta[x, y, r]$ are of class C'''.[5] It is also supposed that $\mathfrak{R}_1$ is an open region of $(2n + 2)$-dimensional points $[x_1, y_{i1}, x_2, y_{i2}]$ in which the functions g, ψ_μ are of class C'', and in which the matrix $(\psi_{\mu x_1}\ \psi_{\mu y_{i1}}\ \psi_{\mu x_2}\ \psi_{\mu y_{i2}})$ has rank p. The terms *differentially admissible*, *terminally admissible* and *admissible* will be employed in the sense used by Morse [15]. This usage of *admissible* is different from that of Bliss [2], [3] and [4]; it seems, however, to be in agreement with that of Bolza ([5], p. 15 and p. 543).

The above problem of Bolza is said to have separated end-conditions if the relations (2.3) fall into two sets

$$\psi_\rho[x_1, y(x_1)] = 0,\ \psi_\sigma[x_2, y(x_2)] = 0 \qquad (\rho = 1, \cdots, k;\ \sigma = k + 1, \cdots, p),$$

and the function g is of the form $g = g_1[x_1, y(x_1)] - g_2[x_2, y(x_2)]$.

---

[5] In order to carry through the analysis of this paper it is actually sufficient to assume that the functions f, φ_β are of class C'' in $\mathfrak{R}$. When only this weaker assumption is made the accessory equations (2.12) may not be expressed as linear equations in η_i', η_i, μ_β, but can be written in the canonical form (2.13).

For a discussion of necessary conditions the reader is referred to the papers of Bliss, Morse, Graves and Hestenes listed in the bibliography. We shall here state explicitly, however, certain relations and conditions which will be used in the present paper.

Let $F[x, y, r, \lambda] = \lambda_0 f[x, y, r] + \lambda_\beta \varphi_\beta[x, y, r]$. An extremal is defined as a differentially admissible arc $y_i(x)$ of class C^2 and a set of multipliers $\lambda_0 =$ constant, $\lambda_\beta(x)$ of class C^1 such that the differential equations $(d/dx)F_{r_i}[x, y, y', \lambda] - F_{y_i}[x, y, y', \lambda] = 0$ are satisfied. As is customary, an extremal will be said to satisfy condition I if

$$(2.5) \qquad [(F - y_i' F_{r_i})\, dx + F_{r_i} dy_i]_1^2 + \lambda_0 dg + e_\mu d\psi_\mu = 0$$

holds for certain constants e_μ and every choice of the differentials dx_1, dy_{i1}, dx_2, dy_{i2}. An extremal is non-singular if the determinant

$$(2.6) \qquad \Delta = \begin{vmatrix} F_{r_i r_j} & \varphi_{\alpha r_i} \\ \varphi_{\beta r_j} & 0_{\beta\alpha} \end{vmatrix}$$

is different from zero at each element $[x, y(x), y'(x), \lambda(x)]$ $(x_1 \leqq x \leqq x_2)$.

In the future we shall be concerned with an extremal of the form $E{:}y_i(x)$, $\lambda_0 = 1$, $\lambda_\beta(x)$ $(x_1 \leqq x \leqq x_2)$. The strengthened Weierstrass condition II_N' is said to hold along this extremal if N is a neighborhood in $[x, y, r, \lambda]$ space of the set $[x, y(x), y'(x), \lambda(x)]$ and

$$(2.7) \quad E[x, y, r, \lambda; \bar{r}] \equiv F[x, y, \bar{r}, \lambda] - F[x, y, r, \lambda] - (\bar{r}_i - r_i)F_{r_i}[x, y, r, \lambda] > 0$$

at every $[x, y, r, \lambda]$ of N and all $(\bar{r}_i)$ such that $[x, y, r]$, $[x, y, \bar{r}]$ are distinct differentially admissible sets. The strengthened Clebsch condition III′ is that at each element $[x, y(x), y'(x), \lambda(x)]$ we have

$$(2.8) \qquad F_{r_i r_j} \pi_i \pi_j > 0$$

for all sets $(\pi_i) \neq (0_i)$ satisfying the equations $\varphi_{\beta r_i} \pi_i = 0$.

Along a given admissible arc E the equations of variation are

$$(2.9) \qquad \Phi_\beta[x, \eta, \eta'] \equiv \varphi_{\beta r_i} \eta_i' + \varphi_{\beta y_i} \eta_i = 0 \qquad (\beta = 1, \cdots, m),$$

$$(2.10) \quad \Psi_\mu[\xi_1, \eta(x_1), \xi_2, \eta(x_2)] \equiv d\psi_\mu[\xi_1, \eta(x_1) + \xi_1 y'(x_1), \xi_2, \eta(x_2) + \xi_2 y'(x_2)] = 0.$$

A set of functions $\eta_i(x)$ of class D^1 satisfying (2.9) on x_1x_2, together with constants ξ_1, ξ_2 satisfying (2.10), is termed an *admissible variation* along E. An admissible arc E satisfies the non-tangency condition if the matrix $(\Psi_{\mu\xi_1}, \Psi_{\mu\xi_2})$ of coefficients of ξ_1, ξ_2 in the expressions Ψ_μ evaluated along E is of rank two.

For an extremal $E{:}y_i(x)$, $\lambda_0 = 1$, $\lambda_\beta(x)$ which satisfies with constants e_μ condition I the second variation is

$$(2.11) \qquad J_2 = 2\gamma[\xi_1, \eta(x_1), \xi_2, \eta(x_2)] + \int_{x_1}^{x_2} 2\omega[x, \eta, \eta']\, dx.$$

In this expression

$$2\omega[x, \eta, \eta'] \equiv F_{r_i r_j} \eta_i' \eta_j' + 2F_{r_i y_j} \eta_i' \eta_j + F_{y_i y_j} \eta_i \eta_j,$$

$$2\gamma[\xi_1, \eta(x_1), \xi_2, \eta(x_2)] \equiv 2H[\xi_1, \eta(x_1) + \xi_1 y'(x_1), \xi_2, \eta(x_2) + \xi_2 y'(x_2)],$$

where as a function of $[dx_1, dy_{i1}, dx_2, dy_{i2}]$,

$$2H[dx_1, dy_{i1}, dx_2, dy_{i2}] \equiv [(F_x - y_i' F_{y_i})\, dx^2 + 2\, F_{y_i} dy_i dx]_1^2 + 2G,$$

and $2G$ is a quadratic form in $[dx_1, dy_{i1}, dx_2, dy_{i2}]$ whose coefficients are the second derivatives of $\mathfrak{g} = g + e_\mu \psi_\mu$. *The explicit form of γ will be needed in the proof of the sufficiency theorem of Section* 6.

Along a non-singular extremal arc the accessory equations are

$$(d/dx)\Omega_{\pi_i}[x, \eta, \eta', \mu] - \Omega_{\eta_i}[x, \eta, \eta', \mu] = 0, \qquad \Phi_\beta[x, \eta, \eta'] = 0, \tag{2.12}$$

where $\Omega[x, \eta, \pi, \mu] \equiv \omega[x, \eta, \pi] + \mu_\beta \Phi_\beta[x, \eta, \pi]$. With the aid of canonical variables x, η_i, ζ_i these equations may be written in canonical form

$$\eta_i' = H_{\zeta_i}, \qquad \zeta_i' = -H_{\eta_i}, \tag{2.13}$$

where $H(x, \eta, \zeta)$ is the Hamiltonian function for the second variation.

An extremal satisfying condition I will be said to also satisfy condition IV_* if along this extremal $J_2 \geqq 0$ for arbitrary non-identically vanishing admissible variations ξ_1, ξ_2, η_i. The strengthened condition obtained by excluding the equality sign in IV_* will be denoted by IV_*'. For a discussion of the relation of IV_* and IV_*' to the various forms of the Mayer (Jacobi) condition which have been used, the reader is referred to Reid ([17], [18], [19]), Hu [11], Hestenes ([9], [10]), Morse [16] and Bliss [4]. For an extremal satisfying III' each of the strengthened forms of the Mayer condition which have been used, with the addition of the non-tangency condition in certain cases,[6] can be proved to imply the above condition IV_*'.

3. **An auxiliary theorem.** The following theorem is a ready consequence of a theorem initially proved by Hestenes ([9], p. 807; see also Bliss [4], p. 112), using a form of the Mayer condition introduced by him. We shall use a slight extension of this theorem which we shall state specifically as Theorem 3.2.

THEOREM 3.1. *Suppose that for a problem of Bolza with separated end-conditions $E: y_i(x), \lambda_0 = 1, \lambda_\beta(x)$ $(x_1 \leqq x \leqq x_2)$ is a non-singular extremal which satisfies with constants e_μ conditions* I *and* IV_*'. *Then there exists a conjugate system of solutions $\eta_i = u_{ik}(x)$, $\zeta_i = v_{ik}(x)$ of the canonical accessory equations* (2.13) *with determinant $| u_{ik}(x) | \neq 0$ on $x_1 x_2$, and such that the inequality*

$$\begin{aligned} &2\Gamma[\xi_1, h_{i1}, \xi_2, h_{i2}] \\ &\quad \equiv 2\gamma[\xi_1, u_{ik}(x_1)h_{k1}, \xi_2, u_{ik}(x_2)h_{k2}] + h_{js} u_{ij}(x_s) v_{ik}(x_s) h_{ks} \Big|_{s=1}^{s=2} > 0 \end{aligned} \tag{3.1}$$

[6] See, in particular, Bliss [4], p. 96 and p. 139.

is valid for all sets $[\xi_1, h_{i1}, \xi_2, h_{i2}] \neq [0, 0_i, 0, 0_i]$ *satisfying*

$$(3.2)\quad L_\mu[\xi_1, h_{i1}, \xi_2, h_{i2}] \equiv \Psi_\mu[\xi_1, u_{ik}(x_1)h_{k1}, \xi_2, u_{ik}(x_2)h_{k2}] = 0 \quad (\mu = 1, \cdots, p).$$

For a problem of Bolza with separated end-conditions, as supposed in the above theorem, conditions (3.2) consist of two sets

$$L_\rho[\xi_1, h_{i1}] = 0, \quad L_\sigma[\xi_2, h_{i2}] = 0 \qquad (\rho = 1, \cdots, k;\ \sigma = k + 1, \cdots, p).$$

Moreover, in this case $2\Gamma[\xi_1, h_{i1}, \xi_2, h_{i2}] = 2\Gamma_1[\xi_1, h_{i1}] + 2\Gamma_2[\xi_2, h_{i2}]$, where Γ_1 and Γ_2 are quadratic forms in their arguments. Consequently, the above conclusion is true if and only if the inequalities $\Gamma_1 > 0$, $\Gamma_2 > 0$ hold for all non-identically vanishing sets $[\xi_1, h_{i1}]$, $[\xi_2, h_{i2}]$ satisfying, respectively, the conditions $L_\rho = 0$, $L_\sigma = 0$. The hypotheses of the above theorem are equivalent to those of Lemma 20.2 of Bliss [4]. The hypotheses of Lemma 7.1 of Hestenes [9] are weaker, however, since in place of IV'_* Hestenes assumed merely the strengthened form of the Mayer condition which he introduced. For the case of fixed end-points a very simple proof of the above theorem is given by Reid ([19], Theorem 4.3).

Under the above hypotheses we have that $E: y_i(x)$, $\lambda_0 = 1$, $\lambda_\beta(x)$ $(x_1 \leqq x \leqq x_2)$ is a non-singular extremal arc, and by the existence theorem for differential equations this extremal is defined for x in a neighborhood of the interval x_1x_2. Since $|u_{ik}(x)| \neq 0$ on x_1x_2, the set $[\xi_1, h_{i1}, \xi_2, h_{i2}]$ is zero if and only if the corresponding set $[\xi_1, u_{ik}(x_1)h_{k1}, \xi_2, u_{ik}(x_2)h_{k2}]$ is zero. If the class of normed sets $[\xi_1, h_{i1}, \xi_2, h_{i2}]$ satisfying

$$L_\mu = 0,\ \xi_1^2 + \xi_2^2 + \|(u_{ik}[x_1]h_{k1})\|^2 + \|(u_{ik}[x_2]h_{k2})\|^2 = 1$$

is non-vacuous, let 2κ denote the minimum of the quadratic form Γ in this class. In view of the above theorem, we have $\kappa > 0$. If this class of normed sets is vacuous, in which case the problem of Bolza reduces to a problem of Lagrange with fixed end-points, κ may be chosen as an arbitrary positive value.

Now suppose that $[X_1, Y_{i1}, X_2, Y_{i2}]$ is in the region $\Re_1$, and such that the extremal $y_i(x)$, $\lambda_0 = 1$, $\lambda_\beta(x)$ of Theorem 3.1 is defined and non-singular on X_1X_2; moreover, that the system of accessory extremals u_{ik}, v_{ik} defined above has determinant $|u_{ik}(x)| \neq 0$ on X_1X_2. For such a set, define $\eta_{is} = Y_{is} - y_i(X_s) = u_{ik}(X_s)h_{ks}$ $(s = 1, 2)$. If $[X_1, Y_{i1}, X_2, Y_{i2}]$ is sufficiently near the end elements $[x_1, y_i(x_1), x_2, y_i(x_2)]$ of E, the set $[x_1(\theta), z_{i1}(\theta), x_2(\theta), z_{i2}(\theta)]$, where $x_s(\theta) = x_s + \theta[X_s - x_s]$, $z_{is}(\theta) = \theta\eta_{is} + y_i[x_s(\theta)]$ $(0 \leqq \theta \leqq 1;\ s = 1, 2)$, belongs to $\Re_1$. In particular, if $[X_1, Y_{i1}, X_2, X_{i2}]$ is terminally admissible we have

$$(3.3)\quad \begin{aligned} 0 &= \psi_\mu[X_1, Y_{i1}, X_2, Y_{i2}] - \psi_\mu[x_1, y_i(x_1), x_2, y_i(x_2)] \\ &= \psi_\mu[X_1, \eta_{i1} + y_i(X_1), X_2, \eta_{i2} + y_i(X_2)] - \psi_\mu[x_1, y_i(x_1), x_2, y_i(x_2)] \\ &= \int_0^1 \{(d/d\theta)\, \psi_\mu[x_1(\theta), z_{i1}(\theta), x_2(\theta), z_{i2}(\theta)]\}\, d\theta \qquad (\mu = 1, \cdots, p). \end{aligned}$$

Now the right-hand members of (3.3) may be written as linear expressions in $[X_1 - x_1, \eta_{i1}, X_2 - x_2, \eta_{i2}]$ whose coefficients are continuous functions of

$[X_1, Y_{i1}, X_2, Y_{i2}]$ and reducing for $[X_1, Y_{i1}, X_2, Y_{i2}] = [x_1, y_i(x_1), x_2, y_i(x_2)]$ to the coefficients of the linear forms Ψ_μ. The following result is then a consequence of Theorem 3.1 and a simple continuity argument.

THEOREM 3.2. *Suppose the hypotheses of Theorem 3.1 are satisfied. Then, using the notation of that theorem, there exists a neighborhood $\mathfrak{N}_0$ of the end-values $[x_1, y_i(x_1), x_2, y_i(x_2)]$ such that if $[X_1, Y_{i1}, X_2, Y_{i2}]$ is in $\mathfrak{N}_0$ the extremal $y_i(x)$, $\lambda_0 = 1$, $\lambda_\beta(x)$ is defined and non-singular on X_1X_2, while $|u_{ik}(x)| \neq 0$ on this interval; moreover, if $[X_1, Y_{i1}, X_2, Y_{i2}]$ is terminally admissible, we have*

$$(3.4)\quad \Gamma[X_1 - x_1, h_{i1}, X_2 - x_2, h_{i2}] \geqq \kappa[(X_1 - x_1)^2 + (X_2 - x_2)^2 + \eta_{i1}\eta_{i1} + \eta_{i2}\eta_{i2}],$$

where $\eta_{is} = Y_{is} - y_i(X_s) = u_{ik}(X_s)h_{ks}$ $(s = 1, 2)$.

That the result of Theorem 3.1 is in general not true for a problem without separated end-conditions is illustrated by the simple problem of minimizing $J[y] = \int_{x_1}^{x_2} y'^2\,dx$ in the class of arcs $y(x)$ $(x_1 \leqq x \leqq x_2)$ satisfying the end-conditions

$$\psi_1[y] \equiv y(x_1) + y(x_2) = 0, \qquad \psi_2 \equiv x_1 = 0, \qquad \psi_3 \equiv x_2 - 1 = 0.$$

Clearly $E{:}y_i(x) \equiv 0$ is the minimizing arc for this problem. Along this arc, $J_2[\eta] \equiv 2J[\eta]$, $\Psi_1 \equiv \psi_1[\eta]$, $\Psi_2 \equiv \xi_1$, $\Psi_3 \equiv \xi_2$, and if the conclusion of Theorem 3.1 were true there would exist a solution u of the accessory equation $u'' = 0$ such that $u(x) \neq 0$ on $(0, 1)$, and $u(1)u'(1)h_2^2 - u(0)u'(0)h_1^2 > 0$ for every set $(h_1, h_2) \neq (0, 0)$ satisfying $u(0)h_1 + u(1)h_2 = 0$. It may be demonstrated readily, however, that there is no such function $u(x)$.

4. **Results concerning the Weierstrass E-function.** This section deals with the behavior of the Weierstrass E-function whenever conditions II'_N and III' are satisfied. As stated in the introduction, the proof here given of conclusion (4.1) of Theorem 4.1 utilizes both of the assumptions II'_N and III'. For the simpler problems of the calculus of variations not involving auxiliary differential equations this conclusion may be proved assuming only II'_N (see Tonelli [20], vol. 1, p. 351). A corresponding proof of this result for the problem of Bolza would be of interest in itself. From the standpoint of the sufficiency theorem, however, assumption of III' is necessary to insure the inequality (4.2).

THEOREM 4.1. *Suppose $E{:}y_i(x)$, $\lambda_0 = 1$, $\lambda_\beta(x)(x_1 \leqq x \leqq x_2)$ is an extremal arc satisfying conditions II'_N, III'. Then there exists in $(2n + m + 1)$ dimensional $[x, y, r, \lambda]$-space a neighborhood N_0 of the values $[x, y(x), y'(x), \lambda(x)]$ $(x_1 \leqq x \leqq x_2)$ belonging to this extremal, and positive constants K, τ_1, τ_2 such that:*

1°. *if $[x, y, r, \lambda]$ is in N_0 and $[x, y, \tilde{r}]$ is a differentially admissible set satisfying $\|\tilde{r} - r\| > K$, then*

$$(4.1)\qquad E[x, y, r, \lambda; \tilde{r}] \geqq \tau_1 \|\tilde{r} - r\|;$$

2°. *if $[x, y, r, \lambda]$ is in N_0, and $[x, y, r]$, $[x, y, \tilde{r}]$ are differentially admissible sets with $\|\tilde{r} - r\| \leqq K$, then*

$$(4.2)\qquad E[x, y, r, \lambda; \tilde{r}] \geqq \tfrac{1}{2}\tau_2 \|\tilde{r} - r\|^2.$$

The proof of conclusion (4.1) here given depends upon the following three lemmas. The neighborhoods which are determined in these lemmas may be supposed, without loss of generality, to be of the form $x_1 - \delta_1 < x < x_2 + \delta_1$, $\| y - y(x) \| < \delta_2$, $\| r - y'(x) \| < \delta_3$, $\| (\lambda_\beta) \| < \delta_4$, where δ_1, δ_2, δ_3, δ_4 are positive constants. Such a neighborhood is convex in the arguments y, r, λ; that is, if $[x, y, r, \lambda]$ and $[x, \bar{y}, \bar{r}, \bar{\lambda}]$ are two elements of such a neighborhood, then the elements $[x, y + \theta(\bar{y} - y), r + \theta(\bar{r} - r), \lambda + \theta(\bar{\lambda} - \lambda)]$ $(0 \leqq \theta \leqq 1)$ are also in the neighborhood.

LEMMA 4.1. *If $E: y_i(x)$, $\lambda_0 = 1$, $\lambda_\beta(x)$ $(x_1 \leqq x \leqq x_2)$ is an extremal satisfying III′, there exists in $[x, y, r, \lambda]$-space a neighborhood N_1 of the elements $[x, y(x), y'(x), \lambda(x)]$ and positive constants c, ϵ_1 such that if $[x, y, r, \lambda]$, $[x, y, \bar{r}, \bar{\lambda}]$ are in N_1 and*

$$(4.3) \qquad | \varphi_{\beta r_i}[x, y, \bar{r}] \pi_i | \leqq \epsilon_1 \| \pi \|,$$

then

$$(4.4) \qquad F_{r_i r_j}[x, y, r, \lambda] \pi_i \pi_j \geqq c \| \pi \|^2.$$

The proof of this lemma is a simple continuity argument, and will be omitted.

LEMMA 4.2. *Suppose $E: y_i(x)$, $\lambda_0 = 1$, $\lambda_\beta(x)$ $(x_1 \leqq x \leqq x_2)$ is an extremal satisfying II'_N. Let N_2 be a second neighborhood of the elements $[x, y(x), y'(x), \lambda(x)]$ with an associated positive constant d such that whenever $[x, y, r, \lambda]$ is in N_2 and $| \delta_\beta | \leqq d$ $(\beta = 1, \cdots, m)$ the element $[x, y, r, \lambda + \delta]$ is in N. Then if $[x, y, r, \lambda]$ is in N_2 and $[x, y, r]$, $[x, y, \tilde{r}]$ are differentially admissible sets, we have*

$$(4.5) \qquad | \varphi_{\beta r_i}[x, y, r] (\tilde{r}_i - r_i) | \leqq (1/d) E[x, y, r, \lambda; \tilde{r}] \quad (\beta = 1, \cdots, m).$$

For since $[x, y, r]$, $[x, y, \tilde{r}]$ are differentially admissible and $[x, y, r, \lambda]$ is in N_2, it is a consequence of II'_N that

$$0 \leqq E[x, y, r, \lambda + \delta; \tilde{r}] = E[x, y, r, \lambda; \tilde{r}] - \delta_\beta \varphi_{\beta r_i}[x, y, r] (\tilde{r}_i - r_i)$$

for all δ_β satisfying $| \delta_\beta | \leqq d$. This inequality is readily seen to imply (4.5).

LEMMA 4.3. *Suppose $E: y_i(x)$, $\lambda_0 = 1$, $\lambda_\beta(x)$ $(x_1 \leqq x \leqq x_2)$ is an extremal satisfying II'_N, III′. Then there exists a positive constant k_1 and a neighborhood N_3 of the elements $[x, y(x), y'(x), \lambda(x)]$ such that whenever $[x, y, r, \lambda]$ is an element of N_3 there are values $z_i = z_i[x, y, r, \lambda]$, $\nu_\beta = \nu_\beta[x, y, r, \lambda]$ satisfying the relations:* (i) *$[x, y, r + z, \lambda + \nu]$ is in N;* (ii) *$\varphi_\beta[x, y, r + z] = 0$ $(\beta = 1, \cdots, m)$;* (iii) *if $[x, y, \tilde{r}]$ is differentially admissible, then*

$$(4.6) \qquad E[x, y, r, \lambda; \tilde{r}] \geqq E[x, y, r + z, \lambda + \nu; \tilde{r}] - k_1 \| z \|^2.$$

Consider the system of equations

$$(4.7) \qquad \begin{aligned} F_{r_i}[x, y, r + z, \lambda + \nu] - F_{r_i}[x, y, r, \lambda] &= 0, \\ \varphi_\beta[x, y, r + z] &= 0. \end{aligned}$$

These equations are satisfied on the closed and bounded set $x_1 \leqq x \leqq x_2$, $y_i = y_i(x)$, $r_i = y'_i(x)$, $\lambda_\beta = \lambda_\beta(x)$, $z_i = 0$, $\nu_\beta = 0$. Moreover, at elements of

this set the Jacobian of the system with respect to the variables z_i, ν_β is the determinant Δ, which is different from zero in view of III′. Hence, by the usual implicit function theorem, there exists a neighborhood N_3 of $[x, y(x), y'(x), \lambda(x)]$ and a constant γ such that for $[x, y, r, \lambda]$ in N_3 the system (4.7) has a unique solution

$$z_i = z_i[x, y, r, \lambda], \nu_\beta = \nu_\beta[x, y, r, \lambda] \tag{4.8}$$

satisfying $z_i z_i + \nu_\beta \nu_\beta < \gamma$, and for such values of $[x, y, r, \lambda]$ the functions z_i, ν_β are of class C^2. We may obviously restrict N_3 so that the functions z_i, ν_β of (4.8) are such that $[x, y, r + z, \lambda + \nu]$ is in N. Under these restrictions the above defined functions z_i, ν_β satisfy (i) and (ii) of the lemma. It is to be remarked that for a simple problem not involving differential equations $\varphi_\beta = 0$ the above lemma is trivial; in this case the solutions (4.8) reduce to $z_i \equiv 0, \nu_\beta \equiv 0$.

Now suppose $[x, y, r, \lambda]$ is in N_3 and $[x, y, \tilde{r}]$ is a differentially admissible set. The relation

$$E[x, y, r, \lambda; \tilde{r}] = E[x, y, r + z, \lambda + \nu; \tilde{r}] + E[x, y, r, \lambda; r + z]$$

is a consequence of (4.7). Moreover,

$$E[x, y, r, \lambda; r + z] = z_i z_j \int_0^1 (1 - \theta) F_{r_i r_j}[x, y, r + \theta z, \lambda]\, d\theta$$
$$\geqq - k_1 \| z \|^2,$$

where k_1 is a positive constant such that if $[x, y, r, \lambda]$ is in N, then

$$| F_{r_i r_j}[x, y, r, \lambda] \pi_i \pi_j | \leqq k_1 \| \pi \|^2 \tag{4.9}$$

for arbitrary sets $\pi = (\pi_i)$. This completes the proof of (iii).

Now suppose that the extremal $E: y_i(x), \lambda_0 = 1, \lambda_\beta(x)$ $(x_1 \leqq x \leqq x_2)$ satisfies both II'_N and III′. Let N_0 be a neighborhood of the elements $[x, y(x), y'(x), \lambda(x)]$ $(x_1 \leqq x \leqq x_2)$ such that $\bar{N}_0$[7] is interior to each of the neighborhood N_1, N_2, N_3, and such that if $[x, y, r, \lambda]$ is in $\bar{N}_0$ and z_i, ν_β are the functions determined in Lemma 4.3, then $[x, y, r + z, \lambda + \nu]$ is in both N_1 and N_2. Since $\bar{N}_0$ is interior to N_1 and N_3 there exists a positive constant d_1 such that whenever $[x, y, r, \lambda]$ is in $\bar{N}_0$ and $u \equiv (u_i)$ has $\| u \| \leqq d_1$, then $[x, y, r + u, \lambda]$ is in both N_1 and N_3. Finally, denote by σ_1 and σ_2 positive constants such that if $[x, y, r, \lambda]$ is in N_3 and z_i are the corresponding functions determined by (4.8) we have $\| z \| \leqq \sigma_1$, $| \varphi_{\beta r_i}[x, y, r + z] z_i | \leqq \sigma_2$. We shall now show that conclusion (4.1) is true if $\tau_1 = \mathrm{Min}\, \{\epsilon_1 d/2; d_1 c/4\}$ and $K = \mathrm{Max}\, \{2d_1; 4k_1\sigma_1^2/cd_1; (2k_1\sigma_1^2/d\epsilon_1) + (2\sigma_2/\epsilon_1)\}$. The constants ϵ_1, c, d, k_1 are those of the above lemmas.

If this conclusion is not true there exists a particular element $[x, y, r, \lambda]$ of N_0 and a set $(\tilde{r}_i)$ such that $[x, y, \tilde{r}]$ is differentially admissible, $\| \tilde{r} - r \| > K$, and $E[x, y, r, \lambda; \tilde{r}] < \tau_1 \| \tilde{r} - r \|$. In view of (4.6) we then have $E[x, y, r + z,$

[7] $\bar{N}_0$ is used to denote N_0, together with its boundary points.

$\lambda + \nu; \tilde{r}] < \tau_1 \| \tilde{r} - r \| + k_1\sigma_1^2$. Since $[x, y, r + z, \lambda + \nu]$ is in N_2 and $\varphi_\beta[x, y, r + z] = 0$, it follows from Lemma 4.2 that

$$| \varphi_{\beta r_i}[x, y, r + z] (\tilde{r}_i - r_i - z_i) | \leqq [\tau_1 \| \tilde{r} - r \| + k_1\sigma_1^2]/d,$$

and

$$| \varphi_{\beta r_i}[x, y, r + z] (\tilde{r}_i - r_i) | \leqq [\tau_1 \| \tilde{r} - r \| + k_1\sigma_1^2]/d + \sigma_2 \leqq \epsilon_1 \| \tilde{r} - r \|.$$

This last inequality is a consequence of $\tau_1 \leqq \epsilon_1 d/2$ and $\| \tilde{r} - r \| > K \geqq (2k_1\sigma_1^2/d\epsilon_1) + (2\sigma_2/\epsilon_1)$. If θ is the value such that $\theta \| \tilde{r} - r \| = d_1$, then $0 < \theta < \frac{1}{2}$. Moreover, since $[x, y, r + z, \lambda + \nu]$ and $[x, y, r + \theta(\tilde{r} - r), \lambda]$ are both in N_1, we have by the use of the mean value theorem and Lemma 4.1 that each of the quantities $E[x, y, r, \lambda; r + \theta(\tilde{r} - r)]$, $E[x, y, r + \theta(\tilde{r} - r), \lambda; r]$ exceeds $c\theta^2 \| \tilde{r} - r \|^2$.

Let $T[x, y, r, \lambda; \tilde{r}] = F[x, y, r, \lambda] + (\tilde{r}_i - r_i)F_{r_i}[x, y, r, \lambda]$. Then

$$\begin{aligned} T[x, y, r + \theta (\tilde{r} - r), \lambda; r + \theta(\tilde{r} - r)] - T[x, y, r, \lambda; r + \theta(\tilde{r} - r)] \\ = E[x, y, r, \lambda; r + \theta(\tilde{r} - r)] > 0, \end{aligned}$$

$$T[x, y, r + \theta(\tilde{r} - r), \lambda; r] - T[x, y, r, \lambda; r] = - E[x, y, r + \theta(\tilde{r} - r), \lambda; r] < 0.$$

Hence, by continuity, there exists a value σ, $0 < \sigma < \theta$, such that

$$T[x, y, r + \theta(\tilde{r} - r), \lambda; r + \sigma(\tilde{r} - r)] - T[x, y, r, \lambda; r + \sigma(\tilde{r} - r)] = 0.$$

This relation implies

$$\begin{aligned} E[x, y, r, \lambda; \tilde{r}] &= E[x, y, r + \theta(\tilde{r} - r), \lambda; \tilde{r}] \\ &\quad + \frac{1 - \sigma}{\sigma} E[x, y, r + \theta(\tilde{r} - r), \lambda; r]. \end{aligned} \tag{4.10}$$

In view of the relations $\theta \| \tilde{r} - r \| = d_1$, $\varphi_\beta[x, y, \tilde{r}] = 0$, it is a consequence of II_N' and Lemma 4.3 that $E[x, y, r + \theta(\tilde{r} - r), \lambda; \tilde{r}] \geqq -k_1\sigma_1^2$. Since $0 < \sigma < \theta < \frac{1}{2}$, $\| \tilde{r} - r \| > K_1 \geqq 4k_1\sigma_1^2/cd_1$, $\tau_1 < d_1c/4$, it then follows from (4.10) that $E[x, y, r, \lambda; \tilde{r}] \geqq \tau_1 \| \tilde{r} - r \|$. This inequality, however, contradicts the assumption that $E[x, y, r, \lambda; \tilde{r}] < \tau_1 \| \tilde{r} - r \|$. Hence conclusion (4.1) is proved.

In order to prove (4.2), consider the neighborhood N_0 defined above. If $[x, y, r, \lambda]$ is in N_0 with $[x, y, r]$ differentially admissible, and $(\tilde{r}_i)$ an arbitrary set such that $[x, y, \tilde{r}]$ is also differentially admissible, and $\| \tilde{r} - r \| < d_1$, it then follows from the mean value theorem and Lemma 4.1 that $E[x, y, r, \lambda; \tilde{r}] \geqq c \| \tilde{r} - r \|^2$. On the other hand, $E[x, y, r, \lambda; \tilde{r}]/ \| \tilde{r} - r \|^2$ is a positive-valued continous function on the bounded closed set: $[x, y, r, \lambda]$ in $\bar{N}_0$, $\varphi_\beta[x, y, r] = 0$, $\varphi_\beta[x, y, \tilde{r}] = 0$, $d_1 \leqq \| \tilde{r} - r \| \leqq K$. Hence there exists a positive constant τ_2 satisfying conclusion 2° of Theorem 4.1.

It is to be emphasized that conclusion 1° of the above theorem is independent of the assumption that $\varphi_\beta[x, y, r] = 0$, whereas this assumption is one of the conditions under which relation (4.2) is valid. The following simple example

illustrates the fact that in general (4.2) is not true if $[x, y, r]$ fails to be differentially admissible: $n = 2, m = 1, f = r_1^2, \varphi_1 = r_2 - y_1, E: y_1(x) \equiv 0 \equiv y_2(x)$. For this problem $E[x, y, r, \lambda; \tilde{r}] \equiv (\tilde{r}_1 - r_1)^2$.

Now consider the convex function $R[t] = \sqrt{(1 + t^2)} - 1 = t^2/(\sqrt{(1 + t^2)} + 1)$ for $t \geqq 0$. We have $t \geqq R[t]$, $t^2 \geqq 2R[t]$, and the following result is an immediate consequence of the above theorem.

THEOREM 4.2. *Suppose the hypotheses of Theorem* 4.1 *are satisfied. Then, using the notation of that theorem, if* $[x, y, r, \lambda]$ *is in* N_0 *and* $[x, y, r]$, $[x, y, \tilde{r}]$ *are differentially admissible sets we have*

$$E[x, y, r, \lambda; \tilde{r}] \geqq \tau_3 R[\| \tilde{r} - r \|], \tag{4.11}$$

where $\tau_3 = \text{Min}\,\{\tau_1, \tau_2\}$.

The use of the function $R[t]$ was suggested by E. J. McShane. McShane has also communicated to the author an independent proof of Theorem 4.2. Further properties of this function are given in the following lemmas.

LEMMA A. *If* $\delta > 0$, *then* $\delta\,\text{Min}\,\{1, \delta\}\, R[rt] \leqq R[\delta t] \leqq \delta\,\text{Max}\,\{1, \delta\}\, R[t]$.

These inequalities follow from the relation $R[\delta t]/R[t] = \delta^2[\sqrt{(1 + t^2)} + 1)/(\sqrt{(1 + \delta^2 t^2)} + 1)]$, and the fact that the quantity in square brackets is between 1 and $1/\delta$.

LEMMA B. *If* $t_2 \geqq t_1 \geqq 0$, *then* $R[t_1 + t_2] \leqq 4\, R[t_2] \leqq 4(R[t_1] + R[t_2])$.

This is a consequence of the relation $R[t_1 + t_2] \leqq R[2t_2]$ and Lemma A.

LEMMA C. *If* $\delta > 0$, *then* $R[t]\,\text{Min}\,\{2, \delta\} \leqq t\,\text{Min}\,\{\delta, t\} \leqq (\sqrt{(1 + \delta^2)} + 1)\, R[t]$.

For consider the function $\rho(t) = t\,\text{Min}\,\{\delta, t\}/R[t] = \text{Min}\,\{\delta, t\}\,(\sqrt{(1 + t^2)} + 1)/t$ for $t > 0$. It is seen that $\rho(t)$ increases on $(0, \delta)$ and decreases on (δ, ∞), and hence $\rho(t) \leqq \rho(\delta) = \sqrt{(1 + \delta^2)} + 1$. Moreover, $\lim_{t\to 0} \rho(t) = 2$, $\lim_{t\to\infty} \rho(t) = \delta$, and therefore $\rho(t) \geqq \text{Min}\,\{2, \delta\}$. Lemma A follows directly from these two inequalities on $\rho(t)$.

We shall now prove the following result, which will be used explicitly in the proof of Theorem 6.1.

THEOREM 4.3. *Suppose that* $E: y_i(x)$, $\lambda_0 = 1$, $\lambda_\beta(x)$ $(x_1 \leqq x \leqq x_2)$ *is an extremal satisfying conditions* II_N', III′, *and that there exists a family of mutually conjugate accessory extremals* $\eta_i = u_{ik}(x)$, $\zeta_i = v_{ik}(x)$ *along this extremal with* $| u_{ik}(x) | \neq 0$ *on* x_1x_2. *For arbitrary sets* (h_i) *let* $y_i(x, h) = y_i(x) + u_{ik}(x)h_k$, $r_i(x, h) = y_i'(x) + u_{ik}'(x)h_k$, $\lambda_\beta(x, h) = \lambda_\beta(x) + \mu_{\beta k}(x)h_k$, *where* $\mu_\beta = \mu_{\beta k}(x)$ *are the multipliers for the accessory extremals* u_{ik}, v_{ik}. *Then for every* $\epsilon > 0$ *there exists a neighborhood* $\mathfrak{F}_{1\epsilon}$ *in xy-space of the* $[x, y(x)]$ *on* E *such that if* $[x, y(x, h)]$ *is in* $\mathfrak{F}_{1\epsilon}$ *and* $[x, y[x, h], \tilde{r}]$ *is differentially admissible, we have*

$$E[x, y(x, h), r(x, h), \lambda(x, h); \tilde{r}] \geqq \tau R[\| \tilde{r} - r(x, h) \|] - \epsilon \| h \|^2, \tag{4.12}$$

where $\tau = \tau_3/4$.

In order to prove (4.12), consider the functions $z_i(x, h) = z_i[x, y(x, h), r(x, h), \lambda(x, h)]$, $\nu_\beta(x, h) = \nu_\beta[x, y(x, h), r(x, h), \lambda(x, h)]$, where $z_i[x, y, r, \lambda]$, $\nu_\beta[x, y, r, \lambda]$ are the functions (4.8). Now $z_i(x, 0) = 0 = \nu_\beta(x, 0)$, and since $\eta_i = u_{ik}(x)$

satisfies the equations $\Phi_\beta[x, \eta, \eta'] = 0$ it follows readily that $z_{ih_j}(x, 0) = 0 = \nu_{\beta h_j}(x, 0)$. Hence corresponding to a given $\epsilon > 0$ there exists a neighborhood $\mathfrak{F}_{1\epsilon}$ in xy-space of E such that whenever $[x, y(x, h)]$ is in $\mathfrak{F}_{1\epsilon}$ we have: (i) $[x, y(x, h), r(x, h), \lambda(x, h)]$ and $[x, y(x, h), r(x, h) + z(x, h), \lambda(x, h) + \nu(x, h)]$ are elements of N_0; (ii) $\| z(x, h) \| \leqq \epsilon_2 \| h \|$, where $\epsilon_2^2 = \epsilon/(2\tau_3 + k_1)$. If $[x, y(x, h)]$ is in $\mathfrak{F}_{1\epsilon}$, then by Lemma 4.3 and Theorem 4.2 we have

$$E[x, y(x, h), r(x, h), \lambda(x, h); \bar{r}] \geqq \tau_3 R[\| \bar{r} - r(x, h) - z(x, h) \|] - \epsilon_2^2 k_1 \| h \|^2. \tag{4.13}$$

In case $\| z(x, h) \| \leqq \| \bar{r} - r(x, h) \| /2$, then $R[\| \bar{r} - r(x, h) - z(x, h) \|] \geqq R[\| \bar{r} - r(x, h) \|/2] \geqq (1/4) R[\| \bar{r} - r(x, h) \|]$, the last inequality by Lemma A. Inequality (4.12) is then true since $\tau = \tau_3/4$, $\epsilon_2^2 < \epsilon/k_1$. If $\| \bar{r} - r(x, h) \| \leqq 2 \| z(x, h) \|$, then $\| \bar{r} - r(x, h) \| \leqq 2\epsilon_2 \| h \|$, and (4.12) is a consequence of (4.13) and the relations

$$\begin{aligned} R[\| \bar{r} - r(x, h) - z(x, h) \|] \geqq 0 &= R[\| \bar{r} - r(x, h) \|] - R[\| \bar{r} - r(x, h) \|] \\ &\geqq R[\| \bar{r} - r(x, h) \|] - (\tfrac{1}{2}) \| \bar{r} - r(x, h) \|^2 \\ &\geqq R[\| \bar{r} - r(x, h) \|] - 2\epsilon_2^2 \| h \|^2. \end{aligned}$$

5. **Elementary integral inequalities.** The integral inequalities of this section are due to McShane, and replace in the proof of Theorem 6.1 certain inequalities previously used by the author. The proof of this theorem originally given by the author made use of integral inequalities similar to those used by Levi and involved Lebesgue integrals. Theorem 5.1 is proved for the general case of Lebesgue integrals. In the application to the proof of Theorem 6.1, however, the functions involved are Riemann integrable.

THEOREM 5.1. *If $h_i(x)$ $(i = 1, \cdots, n)$ are absolutely continuous on $a \leqq x \leqq b$, and $\| h(x) \| \leqq \delta$ on this interval, then*

$$\int_a^b \| h \| \cdot \| h' \| \, dx \leqq d_1 \left[\int_a^b R[\| h' \|] \, dx + \| h(a) \|^2 \right], \tag{5.1}$$

$$\int_a^b \| h \|^2 \, dx \leqq d_2 \left[\int_a^b R[\| h' \|] \, dx + \| h(a) \|^2 \right], \tag{5.2}$$

where $d_1 = 4(\sqrt{(1 + \delta^2)} + 1) \operatorname{Max} \{1, b - a\}$, $d_2 = 3(b - a)d_1$.

If we define

$$H(x) = \| h(a) \| + \int_a^x \| h'(t) \| \, dt, \tag{5.3}$$

then for every x on ab we have

$$\| h(x) - h(a) \| = \left\| \int_a^x h'(t) \, dx \right\| \leqq \int_a^x \| h'(t) \| \, dt^{8} \tag{5.4}$$

[8] Inequality (5.4) follows from Jensen's inequality since $\| u \|$ is a convex function of the vector u. It may also be proved by the use of Cauchy's inequality $| u_i v_i | \leqq \| u \| \cdot \| v \|$.

and hence $H(x) \geqq \| h(a) \| + \| h(x) - h(a) \| \geqq \| h(x) \|$. Therefore,

$$\int_a^b \| h \| \cdot \| h' \| \, dx \leqq \int_a^b \text{Min}\, \{\delta, H(x)\}\, H'(x)\, dx$$

$$\leqq \text{Min}\left\{\int_a^b \delta H'(x)\, dx, \int_a^b H(x)\, H'(x)\, dx\right\}$$

$$\leqq \text{Min}\, \{\delta H(b), H^2(b)\} = H(b)\, \text{Min}\, \{\delta, H(b)\}$$

$$\leqq (\sqrt{(1 + \delta^2)} + 1)\, R[H(b)]. \tag{5.5}$$

The last inequality of (5.5) is by Lemma C.

On the other hand, since $R[t]$ is convex we have by Jensen's inequality and Lemmas A and B that

$$(b - a)^{-1} \int_a^b R[\| h' \|]\, dx$$

$$\geqq R\left[(b - a)^{-1} \int_a^b \| h' \| \, dx\right] = R[(b - a)^{-1}(H(b) - H(a))]$$

$$\geqq (b - a)^{-1}\, \text{Min}\, \{1, (b - a)^{-1}\}\, R[H(b) - H(a)]$$

$$\geqq (b - a)^{-1}\, \text{Min}\, \{1, (b - a)^{-1}\}((1/4)R[H(b)] - R[H(a)]). \tag{5.6}$$

Inequality (5.1) is then obtained by combining (5.5), (5.6), and making use of the relation $R[H(a)] \leqq H^2(a) = \| h(a) \|^2$.

To obtain (5.2), we note that for each x on ab,

$$\| h(x) \|^2 = \| h(a) \|^2 + \int_a^x 2[h_i(t) h_i'(t)]\, dt$$

$$\leqq \| h(a) \|^2 + 2 \int_a^b \| h \| \cdot \| h' \| \, dx,$$

and therefore,

$$\int_a^b \| h \|^2\, dx \leqq (b - a)\left[\| h(a) \|^2 + 2 \int_a^b \| h \| \cdot \| h' \| \, dx\right]. \tag{5.7}$$

Inequality (5.2) is then a consequence of(5.1) and (5.7).

6. **Sufficient conditions for a strong relative minimum.** The results of the preceding sections will now be used to prove the following sufficiency theorem for the general problem of Bolza. This theorem has been proved by Hestenes ([9], Theorem 9.2; see also Bliss [4], Theorem 21.1) by the classical field methods of the calculus of variations.

THEOREM 6.1. *Suppose that* $E: y_i(x), \lambda_0 = 1, \lambda_\beta(x)\ (x_1 \leqq x \leqq x_2)$ *is an extremal arc satisfying with constants* e_μ *conditions* I, II_N', III' *and* IV_*'. *Then there exists a neighborhood* $\mathfrak{F}$ *of* E *in xy-space and a neighborhood* $\mathfrak{N}$ *of the ends of* E *in* $[x_1, y_{i1}, x_2, y_{i2}]$*-space such that* $J(C) > J(E)$ *for every admissible arc* C *in* $\mathfrak{F}$ *with ends in* $\mathfrak{N}$ *and not identical with* E.

It has been pointed out by Hestenes [9] (see also Bliss [4], p. 119) that in order to prove this theorem for the general problem of Bolza *it is sufficient to show that the theorem is true for a problem of Bolza with separated end-conditions and satisfying the non-tangency condition. Consequently, in the following proof we shall assume that these further conditions are satisfied.* As indicated by the example at the end of Section 3, the conclusion of Theorem 3.1 is in general not true along an arc E which satisfies for a problem with mixed end-conditions the hypotheses of Theorem 6.1. It is to be remarked, however, that when the conclusion of Theorem 3.1 holds along such an arc the following expansion method gives a sufficiency theorem without the introduction of an equivalent problem with separated end-conditions.

It follows from hypotheses I and III′ that $E: y_i(x)$, $\lambda_0 = 1$, $\lambda_\beta(x)$ $(x_1 \leqq x \leqq x_2)$ is a non-singular extremal arc. Denote by $u_{ik}(x)$, $v_{ik}(x)$ a mutually conjugate family of accessory extremals along E satisfying the conditions of Theorem 3.2. The multipliers corresponding to u_{ik}, v_{ik} are $\mu_\beta = \mu_{\beta k}(x)$. Now suppose that C: $Y_i(x)$ $(X_1 \leqq x \leqq X_2)$ is an admissible arc with end-values $[X_1, Y_i(X_1), X_2, Y_i(X_2)]$ in the $\mathfrak{N}_0$ neighborhood of $[x_1, y_i(x_1), x_2, y_i(x_2)]$ defined by Theorem 3.2. Let $\eta_i(x) = Y_i(x) - y_i(x)$ on X_1X_2, and define functions $h_i(x)$ by the equations $\eta_i(x) = u_{ik}(x)h_k(x)$; finally, let $\mu_\beta(x) = \mu_{\beta k}(x)h_k(x)$. Corresponding to a given arc C of this type we shall set $u_i' = u_{ik}'(x)h_k(x)$, $v_i = u_{ik}(x)h_k'(x)$. We have, in the notation of Section 4, $Y_i(x) = y_i(x, h[x])$, $y_i'(x) + u_i'(x) = r_i(x, h[x])$, $\lambda_\beta(x) + \mu_\beta(x) = \lambda_\beta(x, h[x])$. Finally, let m_0, M_0 denote positive constants such that for every element $[X_1, Y_{i1}, X_2, Y_{i2}]$ in $\mathfrak{N}_0$ we have $m_0 \,||\, h \,||\, \leqq \,||\, (u_{ik}[x]h_k) \,||\, \leqq M_0 \,||\, h \,||$ $(X_1 \leqq x \leqq X_2)$ for arbitrary sets $h = (h_i)$.

If we write, for brevity, $\Delta g = g[X_1, Y(X_1), X_2, Y(X_2)] - g[x_1, y(x_1), x_2, y(x_2)]$, then

$$\begin{aligned} J[C] - J[E] &= \Delta g + \int_{X_1}^{X_2} f[x, Y, Y']\, dx - \int_{x_1}^{x_2} f[x, y, y']\, dx \\ &= \Delta g + \int_{X_1}^{X_2} (F[x, Y, Y', \lambda + \mu] - F[x, y, y', \lambda])\, dx \\ &\quad - \int_{x_1}^{X_1} F[x, y, y', \lambda]\, dx - \int_{X_2}^{x_2} F[x, y, y', \lambda]\, dx \\ &= \Delta g + J^0 - J^1 - J^2. \end{aligned} \tag{6.1}$$

Now it is readily verified that

$$\begin{aligned} J^0 = &\int_{X_1}^{X_2} E[x, Y, y' + u', \lambda + \mu; Y']\, dx \\ &+ \int_{X_1}^{X_2} (F[x, Y, y' + u', \lambda + \mu] - F[x, y, y', \lambda])\, dx \\ &+ \int_{X_1}^{X_2} v_i F_{r_i}[x, Y, y' + u', \lambda + \mu]\, dx. \end{aligned}$$

By the theorem of the mean we have

$$
\begin{aligned}
&F[x, Y, y' + u', \lambda + \mu] - F[x, y, y', \lambda] \\
(6.2) \qquad &\qquad = F_{r_i} u_i' + F_{y_i} \eta_i + \Omega[x, \eta, u', \mu] + B(x), \\
&F_{r_i}[x, Y, y' + u', \lambda + \mu] = F_{r_i} + F_{r_i r_k} u_k' + F_{r_i y_k} \eta_k + F_{r_i \lambda_\beta} \mu_\beta + B_i^*(x),
\end{aligned}
$$

and for every $\epsilon > 0$ there exists a corresponding $d_\epsilon > 0$ such that whenever $\| \eta \| < d_\epsilon$ we have $| B(x) | \leqq \epsilon \| h(x) \|^2$, $\| B^*(x) \| \leqq (\epsilon/M_0) \| h(x) \|$. It is understood that the partial derivatives of F occurring in the right-hand members of (6.2) have the arguments $[x, y(x), y'(x), \lambda(x)]$. It follows from (6.2) that

$$
\begin{aligned}
J^0 = &\int_{X_1}^{X_2} E[x, Y, y' + u', \lambda + \mu; Y'] \, dx + \int_{X_1}^{X_2} (F_{r_i} \eta_i' + F_{y_i} \eta_i) \, dx \\
(6.3) \qquad &+ (1/2) \int_{X_1}^{X_2} (2\Omega[x, \eta, \eta', \mu] - F_{r_i r_j} v_i v_j) \, dx + \int_{X_1}^{X_2} (B + v_i B_i^*) \, dx.
\end{aligned}
$$

In view of the Euler-Lagrange equations the second integral in (6.3) reduces to $F_{r_i}(X_s)\eta_i(X_s) \big|_{s=1}^{s=2}$. By the Clebsch transformation (see Bliss [1]; also [2], p. 739) the third integral in (6.3) is equal to $h_j(X_s)u_{ij}(X_s)v_{ik}(X_s)h_k(X_s) \big|_{s=1}^{s=2}$. Let $\mathfrak{F}_{2\epsilon}$ be a neighborhood of E such that if C is in $\mathfrak{F}_{2\epsilon}$ then $\| \eta(x) \| \leqq d_\epsilon$ on X_1X_2. Consequently, if C is an admissible arc in $\mathfrak{F}_{2\epsilon}$ with end-points in $\mathfrak{N}_0$ we have

$$
\begin{aligned}
J^0 \geqq &\int_{X_1}^{X_2} E[x, Y, y' + u', \lambda + \mu; Y'] \, dx + [F_{r_i} \eta_i + (\tfrac{1}{2}) h_j u_{ij} v_{ik} h_k]_{X_1}^{X_2} \\
(6.4) \qquad &- \epsilon \int_{X_1}^{X_2} (\| h \|^2 + \| h \| \cdot \| h' \|) \, dx.
\end{aligned}
$$

As in Section 2, let $\mathfrak{g}[x_1, y_{i1}, x_2, y_{i2}] = g[x_1, y_{i1}, x_2, y_{i2}] + e_\mu \psi_\mu[x_1, y_{i1}, x_2, y_{i2}]$, where e_μ are the constants with which the end-values of E satisfy the transversality conditions of I. Consider the function

$$
\begin{aligned}
&\mathfrak{g}^*[t_1, w_{i1}, t_2, w_{i2}] = \mathfrak{g}[t_1, w_{i1} + y_i(t_1), t_2, w_{i2} + y_i(t_2)] \\
(6.5) \qquad &\qquad - \int_{x_1}^{t_1} F[x, y, y', \lambda] \, dx - \int_{t_2}^{x_2} F[x, y, y', \lambda] \, dx + F_{r_i}(t_s) w_{is} \big|_{s=1}^{s=2},
\end{aligned}
$$

where the functions y_i, y_i', λ_β in (6.5) are the functions belonging to the extremal $y_i(x)$, $\lambda_0 = 1$, $\lambda_\beta(x)$. Now expand $\mathfrak{g}^*$ by Taylor's formula about the point $t_s = x_s$, $w_{is} = 0$ $(s = 1, 2)$, and evaluate at $t_s = X_s$, $w_{is} = \eta_i(X_s) = Y_i(X_s) - y_i(X_s)$. It is a consequence of the Euler-Lagrange equations and transversality conditions that

$$
\begin{aligned}
(6.6) \qquad \Delta\mathfrak{g}^* &\equiv \mathfrak{g}^*[X_1, \eta_i(X_1), X_2, \eta_i(X_2)] - \mathfrak{g}^*[x_1, 0_i, x_2, 0_i] \\
&= \gamma[X_1 - x_1, \eta_i(X_1), X_2 - x_2, \eta_i(X_2)] + \gamma^*,
\end{aligned}
$$

and for every $\epsilon > 0$ there exists a bounded neighborhood $\mathfrak{N}_\epsilon$ of the end-values of E which is interior to the $\mathfrak{N}_0$ neighborhood of Theorem 3.2 and such that if $[X_1, Y_i(X_1), X_2, Y_i(X_2)]$ is in $\mathfrak{N}_\epsilon$, then

$$
(6.7) \qquad | \gamma^* | < \epsilon[(X_1 - x_1)^2 + (X_2 - x_2)^2 + \| \eta(X_1) \|^2 + \| \eta(X_2) \|^2].
$$

If the arc C is admissible we have $\Delta\mathfrak{g} = \Delta g$, and

$$\Delta\mathfrak{g}^* = \Delta g - J^1 - J^2 + F_{r_i}(X_s)\eta_i(X_s)\,\Big|_{s=1}^{s=2}. \tag{6.8}$$

In view of the continuity of the functions $u_{ik}(x)$, $v_{ik}(x)$ we may also suppose that the neighborhood $\mathfrak{N}_\epsilon$ is so restricted that whenever $[X_1, Y_{i1}, X_2, Y_{i2}]$ is an element of $\mathfrak{N}_\epsilon$ we have

$$| h_j[u_{ij}(X_s)v_{ik}(X_s) - u_{ij}(x_s)v_{ik}(x_s)]h_k | < 2\epsilon \,||\, h \,||^2 \qquad (s = 1, 2) \tag{6.9}$$

for arbitrary sets $h = (h_i)$.

The above relations imply for an admissible arc in the $\mathfrak{F}_{2\epsilon}$ neighborhood of E with end-points in $\mathfrak{N}_\epsilon$ the following inequality:

$$\begin{aligned} J[C] - J[E] \geqq \Gamma[X_1 - x_1, h_i(X_1), X_2 - x_2, h_i(X_2)] \\ + \int_{X_1}^{X_2} E[x, Y, y' + u', \lambda + \mu; Y']\,dx - \epsilon\{(X_1 - x_1)^2 + (X_2 - x_2)^2 \\ + (M_0 + 1)[||\, h(X_1) \,||^2 + ||\, h(X_2) \,||^2] + \int_{X_1}^{X_2} [||\, h \,||^2 + ||\, h \,|| \cdot ||\, h' \,||]\,dx\}. \end{aligned} \tag{6.10}$$

For a given $\epsilon > 0$ let $\mathfrak{F}_\epsilon$ denote a bounded neighborhood of E in xy-space which is interior to both the neighborhood $\mathfrak{F}_{1\epsilon}$ of Theorem 4.3 and the neighborhood $\mathfrak{F}_{2\epsilon}$ defined above. Suppose $C: Y_i(x)$ $(X_1 \leqq x \leqq X_2)$ is an admissible arc in $\mathfrak{F}_\epsilon$ with end-points in $\mathfrak{N}_\epsilon$. Since $\mathfrak{N}_\epsilon$ is interior to the neighborhood $\mathfrak{N}_0$ of Theorem 3.2, inequality (3.4) is applicable to the quadratic form Γ. We also have that $E[x, Y, y' + u', \lambda + \mu; Y']$ satisfies inequality (4.12). The functions $h_i(x)$ determined by C are absolutely continuous, and $||\, h'(x) \,||$, $R[||\, h'(x) \,||]$, and $R[||\, v(x) \,||]$ are Riemann integrable on X_1X_2. Moreover, since $m_0 \,||\, h' \,|| \leqq ||\, v \,||$ we have by Lemma A that $R[||\, h' \,||] \leqq R[||\, v \,||/m_0] \leqq d_3\, R[||\, v \,||]$, where $d_3 = (1/m_0)\text{ Max }\{1, 1/m_0\}$. Combining these inequalities, it follows that for an admissible arc C in $\mathfrak{F}_\epsilon$ with end-points in $\mathfrak{N}_\epsilon$ we have

$$\begin{aligned} J[C] - J[E] \geqq \sum_{s=1}^{2} \Big[(\kappa - \epsilon)(X_s - x_s)^2 + [\kappa m_0 - \epsilon\{M_0 + 1 + 2d_2 + d_1\}] \cdot \\ ||\, h(X_s) \,||^2 \Big] + [\tau - \epsilon d_3(2d_2 + d_1)] \int_{X_1}^{X_2} R[||\, v(x) \,||]\,dx. \end{aligned} \tag{6.11}$$

Now d_1 and d_2 depend upon the particular arc C in that they involve $(X_2 - X_1)$ and a quantity δ such that $||\, h(x) \,|| < \delta$ on X_1X_2. These quantities, however, are uniformly bounded when ϵ is restricted to a bounded set of values, for example $0 < \epsilon \leqq 1$, and C is in the $\mathfrak{F}_\epsilon$ neighborhood of E with end-points in the $\mathfrak{N}_\epsilon$ neighborhood of the end-points of E. Hence there exists an $\epsilon_0 > 0$ such that for $0 < \epsilon < \epsilon_0$ all the quantities $\kappa - \epsilon$, $\kappa m_0 - \epsilon\{M_0 + 1 + 2d_2 + d_1\}$, $\tau - \epsilon d_3(2d_2 + d_1)$ are positive. Let $\mathfrak{F} = \mathfrak{F}_\epsilon$, $\mathfrak{N} = \mathfrak{N}_\epsilon$ with $0 < \epsilon < \epsilon_0$. Then for an

arbitrary admissible arc C in $\mathfrak{F}$ with end-points in $\mathfrak{N}$ we have $J[C] \geqq J[E]$, and the equality sign holds if and only if

$$0 = \int_{X_1}^{X_2} R[\|\, v(x)\, \|]\, dx = X_s - x_s = \|\, h(X_s)\, \| \qquad (s = 1, 2),$$

that is, if and only if C is identical with E. We have established, therefore, the conclusion of Theorem 6.1.

As emphasized above, in the proof of Theorem 6.1 all the integrals involved may be taken as Riemann integrals. Now let $\mathfrak{F}$ and $\mathfrak{N}$ denote the neighborhoods determined above, and suppose that $C: Y_i(x)$ $(X_1 \leqq x \leqq X_2)$ is an arc in $\mathfrak{F}$ whose end-values are in $\mathfrak{N}$ and terminally admissible, while $Y_i(x)$ $(i = 1, \cdots, n)$ are absolutely continuous and satisfy $\varphi_\beta[x, Y, Y'] = 0$ almost everywhere on X_1X_2. Interpreting the integrals as Lebesgue integrals we have as before that $J[C] \geqq J[E]$ and the equality sign holds if and only if C is identical with E. It is to be remarked that for such arcs C the integrals of the functions $f[x, Y, Y']$ and $E[x, Y, y' + u', \lambda + \mu; Y']$ may be equal to $+\infty$.

If in the proof of Theorem 6.1 use is made of integral inequalities similar to those introduced by Levi, instead of the inequalities of Section 5 suggested by McShane, then the extension of this theorem corresponding to that of the preceding paragraph is less general. It is then necessary to assume that the functions defining C are such that $\|\, Y'(x)\, \|^2$ is Lebesgue integrable.

It is readily seen that the above method gives an expansion proof of Lindeberg's theorem for a semi-strong relative minimum. Conclusion 1° of Theorem 4.1 is not involved in this proof. Such an expansion proof of a semi-strong relative minimum has been given by Levi [12] for the plane problem with fixed end points.

BIBLIOGRAPHY

1. Bliss, *The transformation of Clebsch in the calculus of variations*, Proceedings of the International Mathematical Congress held in Toronto, vol. 1 (1924), pp. 589–603.
2. Bliss, *The problem of Lagrange in the calculus of variations*, American Journal of Mathematics, vol. 52 (1930), pp. 673–744.
3. Bliss, *The problem of Bolza in the calculus of variations*, Annals of Mathematics, vol. 33 (1932), pp. 261–274.
4. Bliss, *The problem of Bolza in the calculus of variations*, Mimeographed notes of lectures delivered at the University of Chicago, Winter Quarter of 1935.
5. Bolza, *Vorlesungen über Variationsrechnung*, Teubner, 1909.
6. Duren, *The development of sufficient conditions in the calculus of variations*, Dissertation (Chicago, 1930), published in Contributions to the Calculus of Variations, 1930, The University of Chicago.
7. Graves, *On the Weierstrass condition for the problem of Bolza in the calculus of variations*, Annals of Mathematics, vol. 33 (1932), pp. 747–752.
8. Hahn, *Über die hinreichenden Bedingungen für ein starkes Extremum beim einfachsten Probleme der Variationsrechnung*, Rendiconti del Circolo Matematico di Palermo, vol. 36 (1913), pp. 379–386.
9. Hestenes, *Sufficient conditions for the problem of Bolza in the calculus of variations*, Transactions of the American Mathematical Society, vol. 36 (1934), pp. 793–818.

10. HESTENES, *On sufficient conditions in the problems of Lagrange and Bolza*, Annals of Mathematics, vol. 37 (1936), pp. 543–551.
11. HU, *The problem of Bolza and its accessory boundary value problem*, Dissertation (Chicago, 1932), published in Contributions to the Calculus of Variations, 1931–1932, The University of Chicago.
12. LEVI, *Sulle condizioni sufficienti per il minimo nel calcolo delle variazioni (integrali forma non parametrica)*, Atti della R. Accademia dei Lincei, (5), vol. 20^2 (1911), pp. 425–431, and pp. 466–469.
13. LEVI, *Sui criterii sufficienti per il massimo e per il minimo nel calcolo delle variazioni*, Annali di Matematica Pura ed Applicata (3), vol. 21 (1913), pp. 173–218.
14. MALNATE, *Sui criterii sufficienti per il massimo e minimo nel calcolo delle variazioni*, Giornale di Matematiche di Battaglini, vol. 57 (1919), pp. 79–102.
15. MORSE, *Sufficient conditions in the problem of Lagrange with variable end-points*, American Journal of Mathematics, vol. 53 (1931), pp. 517–546.
16. MORSE, *Sufficient conditions for the problem of Bolza in the calculus of variations*, Transactions of the American Mathematical Society, vol. 37 (1935), pp. 147–160.
17. REID, *A boundary value problem associated with the calculus of variations*, American Journal of Mathematics, vol. 54 (1932), pp. 769–790.
18. REID, *Analogues of the Jacobi condition for the problem of Mayer in the calculus of variations*, Annals of Mathematics, vol. 35 (1934), pp. 836–848.
19. REID, *The theory of the second variation for the nonparametric problem of Bolza*, American Journal of Mathematics, vol. 57 (1935), pp. 573–586.
20. TONELLI, *Fondamenti di calcolo delle variazioni*, vols. 1 and 2, Bologna, 1921.

UNIVERSITY OF CHICAGO
CHICAGO, ILLINOIS

Annals of Mathematics
Vol. 38, No. 3, July, 1937

PLATEAU'S PROBLEM AND DIRICHLET'S PRINCIPLE

By Richard Courant

(Received December 15, 1936; revised May 11, 1937)

Contents

1. Introduction.
2. Preliminary Lemmas on the Dirichlet Integral.
3. Plateau's Problem for one contour ($k = 1$).
4. Plateau's Problem for k contours. Preparations.
5. Solution of Plateau's Problem for $k = 2$.
6. Plateau's Problem for arbitrary k.
7. Solution of Plateau's Problem and the problem of least area based on conformal mapping.
8. Plateau's Problem for one-sided minimal surfaces and for minimal surfaces of higher genus.

1. Introduction

The construction of a surface of least area bounded by a given contour Γ in the 3-dimensional Euclidean x_1, x_2, x_3-space is one of the classical problems in the calculus of variations.[1] To formulate the problem analytically we suppose the surface under consideration represented by functions $x_\mu(u, v)$ of two parameters u, v (or by a vector $\mathfrak{x}(u, v)$ with the x_μ as components) in a given domain B of the u, v-plane with the boundary C; these functions shall be continuous in the closed domain $B + C$, have piece-wise continuous second derivatives in B and map C on Γ. Then the problem requires us to minimize by one of these admissible functions the integral

$$A(\mathfrak{x}) = \iint_B \sqrt{(EG - F^2)}\, du\,dv, \tag{1}$$

where we have with the usual notations

$$E = \mathfrak{x}_u^2 = \sum_\mu \left(\frac{\partial x_\mu}{\partial u}\right)^2; \qquad G = \mathfrak{x}_v^2 = \sum_\mu \left(\frac{\partial x_\mu}{\partial v}\right)^2;$$

$$F = \mathfrak{x}_u \mathfrak{x}_v = \sum_\mu \frac{\partial x_\mu}{\partial u} \cdot \frac{\partial x_\mu}{\partial v}.$$

This integral $A(\mathfrak{x})$ is invariant under arbitrary transformations of the parameters u, v and their domain B, which, if Γ is a simple Jordan curve, may be chosen as the unit circle $u^2 + v^2 < 1$.

[1] Cf. throughout this paper as a reference the excellent report by Radò "The problem of Plateau," Ergebnisse der Mathematik, II, 2, Berlin, 1933.

The Euler conditions of this variational problem form a system of non-linear partial differential equations in B with additional non-linear conditions, expressing not only that the required surface is bounded by Γ, but also that it has mean curvature zero or is a "*minimal surface.*" Replacing the problem of least area by that of a minimal surface bounded by Γ—this problem is called Plateau's problem[2]—one has to meet first the difficulty consisting in the non-linearity of the differential equations and second the difficulty consisting in the non-linearity of the additional conditions. Riemann and after him Weierstrass, H. A. Schwarz, Darboux and others have already entirely disposed of the first difficulty. By taking advantage of the freedom in the choice of the parameters one can linearize the differential equations of the least area problem. Suppose it is possible to introduce isometric parameters u, v on the surfaces S considered in our variational problem, i.e. parameters for which

$$E - G = 0; \qquad F = 0, \tag{2}$$

or, in other words, parameters which correspond to a conformal mapping of the surface S on the u, v-plane. Then we have

$$A(\mathfrak{x}) = D(\mathfrak{x}) = D_B(\mathfrak{x}) \tag{3}$$

with

$$D(\mathfrak{x}) = \tfrac{1}{2}\int\int_B (E + G)\, du\, dv = \tfrac{1}{2}\int\int_B (\mathfrak{x}_u^2 + \mathfrak{x}_v^2)\, du\, dv \tag{4}$$

where $D(\mathfrak{x})$ is the classical Dirichlet integral.

In general, if u and v are not necessarily isometric parameters, we have that $A(\mathfrak{x}) \leqq \int\int_B \sqrt{EG}\; du\, dv$, and therefore

$$A(\mathfrak{x}) \leqq D(\mathfrak{x}), \tag{5}$$

where

$$A(\mathfrak{x}) = D(\mathfrak{x}) \tag{5a}$$

holds if and only if

$$E - G = F = 0. \tag{2}$$

Under the conditions (2) the Euler equations for $A(\mathfrak{x}) = D(\mathfrak{x})$ simply become linear:

$$\Delta x_\mu = 0; \qquad \left(\Delta = \frac{\partial^2}{\partial u^2} + \frac{\partial^2}{\partial v^2}\right) \tag{6}$$

or

$$\Delta \mathfrak{x} = 0 \tag{7}$$

[2] On account of the famous experiments by Plateau. Cf. Radò, loc. cit.

This leads to the definition of a minimal surface as one for which, in a suitable parametric representation, we have the linear differential equation

$$\Delta \mathfrak{x} = 0 \tag{7}$$

with the additional non-linear conditions

$$E - G = F = 0. \tag{8}$$

In other words, Plateau's problem requires us to find a *"potential vector"* $\mathfrak{x}(u, v)$ in $B + C$ which gives a conformal mapping of B on a surface bounded by Γ.[3]

If we consider potential functions x_μ as the real parts $\Re f_\mu(u + iv) = \Re f_\mu(w)$ of analytic functions $f_\mu(w)$ of the complex variable $w = u + iv$, it follows at once that

$$(E - G) - 2iF = \sum_\mu [f'_\mu(w)]^2 = \varphi(w) \tag{8'}$$

is an analytic function of $w = u + iv$ for any potential vector $\mathfrak{x}$. The condition (2) that our potential surface S with $\Delta\mathfrak{x} = 0$ be a minimal surface can be expressed by

$$\varphi(w) = 0. \tag{9}$$

It may be remarked that the boundary conditions represent two non-linear relations between x_1, x_2, x_3 on C while the two additional relations (2) or (9), which have the appearance of two additional non-linear partial differential equations, amount to only one more requirement of the type of a boundary condition. For, since $\varphi(w)$ is an analytic function with *any* potential vector $\mathfrak{x}$, this function must vanish identically if its real part $E - G$ has the boundary values zero and if in addition the imaginary part vanishes at one point.

On this definition of a minimal surface Riemann, Schwarz, and others have based the solution of Plateau's Problem for many interesting particular cases—thereby deviating from the calculus of variations as the starting point. More recently Garnier obtained a rather general result by pursuing the line of Riemann's method. But it was only by combining Riemann's idea again with the viewpoint and with the method of the calculus of variations that eventually, in 1932, T. Radò and J. Douglas independently succeeded in giving a satisfactorily general existence proof.[4]

---

[3] This linearization of the non-linear Euler equations corresponds exactly to that of the Euler equations of geodesic lines by choosing the arc length as parameter. Incidentally the linearization of the differential equations of a minimal surface is only a special case of a much more general fact concerning quasi-linear partial differential equations of the second order in two independent variables (cf. Courant-Hilbert, *Methoden der mathematischen Physik*, vol. II, Chapter III, in press).

[4] Quite a different approach to the problem of least area from that of the calculus of variations was attempted by Lebesgue. More recently McShane has also pursued this line, and has established the solution for $k = 1$ under remarkably general conditions for the surfaces in competition. Cf. Radò's report.

The merit of Douglas's works extends beyond the solution of the original problem of Plateau. Apart from a somewhat greater generality with respect to admissible boundaries Γ and apart from his remark that the method applies to any number m of dimensions of the $x_1, \cdots, x_m$-space he has attacked the problem itself on a broader front. He envisaged the more general and much more difficult task of the construction of a minimal surface which is bounded by k given contours (Jordan curves) $\Gamma_1, \cdots, \Gamma_k$ and which has a prescribed topological structure, e.g. is required to be one-sided or two-sided and to have a prescribed genus. Douglas has so far published a solution of the problem for two-sided minimal surfaces of the genus zero for $k = 1$ and $k = 2$ and also for one-sided surfaces of the type of a Moebius strip with $k = 1$. Moreover he has announced the publication of a solution in the general case.[5]

Radò achieved his goal by first approximating to the lower bound of the integral $A(\mathfrak{x})$ by means of polyhedral surfaces and then by mapping these surfaces conformally on the unit circle. Douglas, on the contrary, lays much emphasis on avoiding the use of conformal mapping and rather on including Riemann's mapping theorem as a consequence of the solution of Plateau's problem for $m = 2$; i.e. for a two-dimensional $\mathfrak{x}$-space.[6]

The representation of the minimal surface by a potential vector $\mathfrak{x}$ and the consideration of the relationships (5) (5a) between $A(\mathfrak{x})$ and $D(\mathfrak{x})$ makes it plausible that $\mathfrak{x}$ is the solution of the *variational problem*: To minimize the Dirichlet integral $D(\mathfrak{x})$ under the condition that $\mathfrak{x}$ maps the circle B on a surface S bounded by Γ.[7] This problem will be called Problem I.

However, Douglas does not make such a variational problem of the classical Riemann-Dirichlet type the basis of his reasoning. Instead, from the beginning, he substitutes in the Dirichlet integral for $\mathfrak{x}$ potential vectors solely and then

---

[5] Journal of Math. and Phys., vol. XV (1936) p. 55-64 and p. 106-123. The second of these papers gives more detailed information about the proof, which is based upon the theory of Riemann's multiply-periodic ϑ-functions on a Riemann-surface.

A complete reference to Douglas' previous papers is included in his article "The problem of Plateau," Bull. Amer. Math. Soc. (1933) p. 227-251.

[6] However, Douglas also applies the theory of conformal mapping of polyhedral surfaces to show that his solution gives the least area.

[7] If we assume the possibility of a conformal mapping on the unit circle for all surfaces admitted to competition in the original variational problem $A(\mathfrak{x}) = \min.$, then from our inequalities (5) and (5a) it follows immediately that the lower limits of $A(\mathfrak{x})$ and $D(\mathfrak{x})$ must be identical. Therefore, the solution of the problem for the Dirichlet integral also solves the original problem for the area and satisfies $E - G = F = 0$ in addition to $\Delta\mathfrak{x} = 0$, because for the solution $\mathfrak{x}$ we have $A(\mathfrak{x}) = D(\mathfrak{x})$. This reasoning for $k = 1$ which plays an important rôle in Radò's proof was later emphasized also by Douglas "The mapping theorem of Koebe and the problem of Plateau," Journ. of Math. and Phys. vol. X (1931) pp. 106-130. It is true that the initial assumption requires some discussion. But it can be verified in a rather elementary manner and with sufficient generality, even in the higher cases for $k > 1$.

transforms the integral $D(\mathfrak{x})$ by means of Poisson's formula into his well-known functional, which contains only boundary values of $\mathfrak{x}$ on C:

$$H(\mathfrak{x}) = \frac{1}{4\pi}\int_0^{2\pi}\int_0^{2\pi} \frac{[\mathfrak{x}(\alpha) - \mathfrak{x}(\beta)]^2}{4\sin^2 \frac{1}{2}(\alpha - \beta)}\, d\alpha\, d\beta, \tag{10}$$

where $\mathfrak{x}(\vartheta)$ are the boundary values of $\mathfrak{x}$ on C as functions of the angle ϑ. Now he starts with the problem of minimizing $H(\mathfrak{x})$ if all those vectors $\mathfrak{x}$ on C which map C in a continuous monotonic way on Γ are admitted to competition.

This ingenious departure from classical lines to a variational problem not implying derivatives makes it easy to establish the existence of a minimizing $\mathfrak{x}$. Thereby the complications in the method are shifted to the task of excluding harmful singularities from the solution and of identifying the solution with the boundary values of a potential vector satisfying (2).

In the case of two contours where the generalization of the boundary functional $H(\mathfrak{x})$ becomes less elementary, these complications are more marked; and this certainly is true all the more in the case of more contours or of higher genus, where Douglas, according to his announcement, will make essential use of the theory of Abelian functions on Riemann surfaces of arbitrary genus, also considering their dependence on the moduli of the surface.

To link Plateau's problem with these deep and beautiful theories will be, when presented in detail, an achievement of highest interest. But it seems worthwhile to avoid the complications arising from the explicit expression by a boundary-functional, and rather to start directly with a Dirichlet Problem I as above, where the vectors $\mathfrak{x}$ in competition need not be potential vectors.[8]

The method which I thus have developed on the line of Dirichlet's Principle not only solves Plateau's problem and the most general problem formulated by Douglas for any number of contours and any prescribed topological type, in a very simple way, but it also allows us to solve Plateau's problem in cases apparently not accessible to Douglas' original method, in which parts of the boundaries are free on prescribed manifolds of any dimension less than m. Moreover, other problems in geometry and hydrodynamics, not as yet solved, seem accessible to this method.

The method consists essentially of two parts. In the first part the variational Problem I is solved. According to the fundamental Dirichlet Principle the solution must be a potential vector. In the second part the solution is shown, by variational methods, to satisfy the condition (9) characterizing a minimal

[8] Radò (l.c.p. 86 ff) has already observed that in the case of one contour, by a reasoning based on such a variational problem, an alternative of Douglas' proof of the relations $E - G = F = 0$ can be given. Moreover, as also was noticed by Radò (l.c.p. 77), in excluding the possibility of a degeneration from his solution, Douglas uses an argument equivalent to a reasoning which has been applied in some of my former papers on the Dirichlet principle and conformal mapping. In the present paper this reasoning is of basic importance for the construction of the solution itself. Cf. Lemmas 5 and 6.

The merit of Douglas's works extends beyond the solution of the original problem of Plateau. Apart from a somewhat greater generality with respect to admissible boundaries Γ and apart from his remark that the method applies to any number m of dimensions of the $x_1, \cdots, x_m$-space he has attacked the problem itself on a broader front. He envisaged the more general and much more difficult task of the construction of a minimal surface which is bounded by k given contours (Jordan curves) $\Gamma_1, \cdots, \Gamma_k$ and which has a prescribed topological structure, e.g. is required to be one-sided or two-sided and to have a prescribed genus. Douglas has so far published a solution of the problem for two-sided minimal surfaces of the genus zero for $k = 1$ and $k = 2$ and also for one-sided surfaces of the type of a Moebius strip with $k = 1$. Moreover he has announced the publication of a solution in the general case.[5]

Radò achieved his goal by first approximating to the lower bound of the integral $A(\mathfrak{x})$ by means of polyhedral surfaces and then by mapping these surfaces conformally on the unit circle. Douglas, on the contrary, lays much emphasis on avoiding the use of conformal mapping and rather on including Riemann's mapping theorem as a consequence of the solution of Plateau's problem for $m = 2$; i.e. for a two-dimensional $\mathfrak{x}$-space.[6]

The representation of the minimal surface by a potential vector $\mathfrak{x}$ and the consideration of the relationships (5) (5a) between $A(\mathfrak{x})$ and $D(\mathfrak{x})$ makes it plausible that $\mathfrak{x}$ is the solution of the *variational problem*: To minimize the Dirichlet integral $D(\mathfrak{x})$ under the condition that $\mathfrak{x}$ maps the circle B on a surface S bounded by Γ.[7] This problem will be called Problem I.

However, Douglas does not make such a variational problem of the classical Riemann-Dirichlet type the basis of his reasoning. Instead, from the beginning, he substitutes in the Dirichlet integral for $\mathfrak{x}$ potential vectors solely and then

---

[5] Journal of Math. and Phys., vol. XV (1936) p. 55–64 and p. 106–123. The second of these papers gives more detailed information about the proof, which is based upon the theory of Riemann's multiply-periodic ϑ-functions on a Riemann-surface.

A complete reference to Douglas' previous papers is included in his article "The problem of Plateau," Bull. Amer. Math. Soc. (1933) p. 227–251.

[6] However, Douglas also applies the theory of conformal mapping of polyhedral surfaces to show that his solution gives the least area.

[7] If we assume the possibility of a conformal mapping on the unit circle for all surfaces admitted to competition in the original variational problem $A(\mathfrak{x}) = \min.$, then from our inequalities (5) and (5a) it follows immediately that the lower limits of $A(\mathfrak{x})$ and $D(\mathfrak{x})$ must be identical. Therefore, the solution of the problem for the Dirichlet integral also solves the original problem for the area and satisfies $E - G = F = 0$ in addition to $\Delta\mathfrak{x} = 0$, because for the solution $\mathfrak{x}$ we have $A(\mathfrak{x}) = D(\mathfrak{x})$. This reasoning for $k = 1$ which plays an important rôle in Radò's proof was later emphasized also by Douglas "The mapping theorem of Koebe and the problem of Plateau," Journ. of Math. and Phys. vol. X (1931) pp. 106–130. It is true that the initial assumption requires some discussion. But it can be verified in a rather elementary manner and with sufficient generality, even in the higher cases for $k > 1$.

transforms the integral $D(\mathfrak{x})$ by means of Poisson's formula into his well-known functional, which contains only boundary values of $\mathfrak{x}$ on C:

$$H(\mathfrak{x}) = \frac{1}{4\pi} \int_0^{2\pi} \int_0^{2\pi} \frac{[\mathfrak{x}(\alpha) - \mathfrak{x}(\beta)]^2}{4 \sin^2 \frac{1}{2}(\alpha - \beta)}\, d\alpha\, d\beta, \tag{10}$$

where $\mathfrak{x}(\vartheta)$ are the boundary values of $\mathfrak{x}$ on C as functions of the angle ϑ. Now he starts with the problem of minimizing $H(\mathfrak{x})$ if all those vectors $\mathfrak{x}$ on C which map C in a continuous monotonic way on Γ are admitted to competition.

This ingenious departure from classical lines to a variational problem not implying derivatives makes it easy to establish the existence of a minimizing $\mathfrak{x}$. Thereby the complications in the method are shifted to the task of excluding harmful singularities from the solution and of identifying the solution with the boundary values of a potential vector satisfying (2).

In the case of two contours where the generalization of the boundary functional $H(\mathfrak{x})$ becomes less elementary, these complications are more marked; and this certainly is true all the more in the case of more contours or of higher genus, where Douglas, according to his announcement, will make essential use of the theory of Abelian functions on Riemann surfaces of arbitrary genus, also considering their dependence on the moduli of the surface.

To link Plateau's problem with these deep and beautiful theories will be, when presented in detail, an achievement of highest interest. But it seems worthwhile to avoid the complications arising from the explicit expression by a boundary-functional, and rather to start directly with a Dirichlet Problem I as above, where the vectors $\mathfrak{x}$ in competition need not be potential vectors.[8]

The method which I thus have developed on the line of Dirichlet's Principle not only solves Plateau's problem and the most general problem formulated by Douglas for any number of contours and any prescribed topological type, in a very simple way, but it also allows us to solve Plateau's problem in cases apparently not accessible to Douglas' original method, in which parts of the boundaries are free on prescribed manifolds of any dimension less than m. Moreover, other problems in geometry and hydrodynamics, not as yet solved, seem accessible to this method.

The method consists essentially of two parts. In the first part the variational Problem I is solved. According to the fundamental Dirichlet Principle the solution must be a potential vector. In the second part the solution is shown, by variational methods, to satisfy the condition (9) characterizing a minimal

[8] Radò (l.c.p. 86 ff) has already observed that in the case of one contour, by a reasoning based on such a variational problem, an alternative of Douglas' proof of the relations $E - G = F = 0$ can be given. Moreover, as also was noticed by Radò (l.c.p. 77), in excluding the possibility of a degeneration from his solution, Douglas uses an argument equivalent to a reasoning which has been applied in some of my former papers on the Dirichlet principle and conformal mapping. In the present paper this reasoning is of basic importance for the construction of the solution itself. Cf. Lemmas 5 and 6.

surface. The domain B of the parameters u, v for $\mathfrak{x}(u, v)$ in all cases can be supposed as bounded by circles, one-sidedness or higher genus being taken care of by proper coordination of certain boundaries. (Also other types of domains B are suitable, and even, in case of higher genus, sometimes preferable.) The variational problem is dealt with, on the basis of elementary lemmas. In the case $k > 1$ or for higher genus etc. some additional conditions are required to prevent degeneration of the domain B.

For the task of the second part two essentially different methods are offered. On the one hand one can avoid any use of theorems on conformal mapping in which case the condition $E - G = F = 0$ appears as a "*natural boundary condition*," arising by suitable variations with respect to the degree of freedom left in the boundary representation, and in higher cases by additional variation with respect to the domain B. For $m = 2$ then, indeed, general theorems on conformal mapping result as special cases.

On the other hand for the second part one can use a quite different method based on simple, but general theorems on conformal mapping of domains in the u, v plane, bounded by Jordan curves. By consenting to the use of those slightly less elementary tools one can simplify the task of the second part considerably.

Radò's and Douglas's methods yield a minimal surface which at the same time has the least possible area under the prescribed conditions. The same is true for the method presented in this paper. However there exist in many cases minimal surfaces with given boundaries which do not furnish the least possible area (e.g. the catenoid, if the endpoints of the generating curve are nearly conjugate). As was first noticed by Max Shiffman[9] such cases of relative minima are also accessible to the method of this paper.

An outline of the method has already been published.[10] In the present paper the problem for k contours and genus zero is treated in full detail with the first method avoiding conformal mapping in §§3, 4, 5, 6. The second variant based on theorems on conformal mapping is given in §7; how the method applies to the case of higher genus and to the case of onesided minimal surfaces is discussed in §8. The proof of a mapping theorem used in §7, the detailed discussion of the topological degenerations of the domains B used in §8, the amplification for the case of free boundaries and other applications and extensions of the method are deferred to subsequent publication.

2. Preliminary Lemmas on the Dirichlet Integral

1. **Boundary value problem and minimum of the Dirichlet integral.** We suppose that we can solve the boundary value problem of the potential equation

[9] Cf. a subsequent publication and footnote 24.

[10] Proceedings of the Natl. Acad. of Sciences, vol. XXII (1936) pp. 367–372, 373–375. Douglas's notes indicating his method for the general case of a prescribed topological structure (see footnote 5) were either published or submitted to the editors while my communications were in proofs. Therefore it seems fitting to state expressly that my publications do not contest Douglas's claim to priority.

$\Delta p(u, v) = 0$ for a domain B in the u, v-plane bounded by k circles $C_1, \cdots, C_k$ or, what is equivalent, that we know Green's function for such a domain. For $k = 1$ the solution is given by the elementary Poisson integral. We further consider as known[11] the following

LEMMA 1. *Let the function $q(u, v)$ range over the set of all functions continuous in the domain B and on its boundary, having piecewise[12] continuous first derivatives in B, and assuming prescribed boundary values on the boundary $C_1, \cdots, C_k$. Let the Dirichlet Integral*

$$D(q) = \iint_B (q_u^2 + q_v^2)\, du\, dv$$

admit finite values. Then the minimal value $d = d(B)$ of $D(q)$ is attained for and only for the function $q = p$, which solves the corresponding boundary value problem of $\Delta p = 0$.

LEMMA 1a. *Let the domain B have an additional circular boundary C^*. Let the function $q(u, v)$ satisfy the same conditions as in Lemma 1 for B and $C_1, \cdots, C_k$, but with no boundary values prescribed on C^*. Then the minimal values d_o of $D(q)$ is attained for and only for the potential function $q = p_o$, which has the prescribed boundary values on $C_1, \cdots, C_k$ and has a vanishing normal derivative on C^*. We have $D(p_o) = d_o = d_o(B)$.*

Here $d_o = d_o(B)$ refers to the minimum with respect to the prescribed boundary values on $C_1, C_2 \cdots C_k$ and, in addition, arbitrary values on C^*.

The function p_o can be obtained as the solution of an ordinary boundary value problem for a circular domain consisting of B and the image of B by reflection on C^*, whereby boundary points corresponding by reflection carry the same boundary values.

We add the remark: The statement subsists if instead of functions q, p vectors $\mathfrak{z}$, $\mathfrak{x}$ are substituted in the Dirichlet Integral.

The following lemma (not used for solving Plateau's Problem in the simplest case $k = 1$ in §3) states that in the variational problem $D(q) = \min = d$ or $D(q) = \min = d_o$ the lower limit d or d_o is not noticeably affected, if we restrict the range of competition for q by imposing the condition that q shall vanish (or be constant) in a point set contained in a sufficiently small neighborhood of a fixed point P of B.

LEMMA 2. *If d is the lower limit of $d(B)$ or $d_o(B)$ as in Lemma 1 or 1a and d_ϵ the same lower limit under the additional condition that q vanishes on a prescribed point*

[11] A simple proof can be found e.g. in Hurwitz-Courant, *Funktionentheorie*, 3rd ed. Berlin, 1930, part 3 and in Courant-Hilbert, *Methoden der mathematischen Physik*, Bd. II, Chapter 7, §1 (in press).

[12] A function is called *piecewise-continuous* in B, if it is continuous except for isolated points and a finite number of arcs of curves with continuous tangents in B where it may have discontinuities of the first kind.

set A inside a circle C_ϵ of radius $r = \epsilon$ around P, then we have $\lim_{\epsilon\to 0} d_\epsilon = d$, *and more precisely* $d \leqq d_\epsilon \leqq d + \sigma(\epsilon)$, *where $\sigma(\epsilon)$ is a quantity depending only on ϵ and tending to zero with ϵ. This relation holds uniformly for all boundary values for which $|q| \leqq M$ with fixed M.*

The same is true if B contains the point at infinity, C_ϵ denotes a circle of radius $1/\epsilon$ around the origin and the point set A lies outside C_ϵ.

Proof. We certainly have $d_\epsilon \geqq d$ because the minimum problem defining d_ϵ originates from that defining d by the addition of new restrictions, thus narrowing the range of competition. We have to show that for sufficiently small ϵ, functions q_ϵ can be found for which the additional condition is satisfied and for which $D(q_\epsilon) < d + \sigma$, where $\sigma = \sigma(\epsilon)$ can be chosen arbitrarily small. Indeed, if p is the function in the original problem for which $D(p) = d$ we obtain from p another function q satisfying the additional conditions by putting $q_\epsilon(u, v) = p\tau$ where in polar coordinates, r being the distance from the point P, we have[13]

$$\tau(r) = 0 \quad \text{for} \quad r \leqq \epsilon$$

$$\tau(r) = 1 \quad \text{for} \quad r \geqq \sqrt{\epsilon}$$

$$\tau(r) = -\frac{2}{\log \epsilon} \log \frac{r}{\epsilon} \quad \text{for} \quad \epsilon \leqq r \leqq \sqrt{\epsilon}.$$

We have further

$$D(\tau) = -\frac{4\pi}{\log \epsilon} \to 0, \quad \text{as} \quad \epsilon \to 0$$

and

$$D(q_\epsilon) \leqq D(p) + D'(p\tau)$$

where the symbol D' denotes that the Dirichlet Integral is to be extended over $r \leqq \sqrt{\epsilon}$. Considering the inequalities $|\tau| \leqq 1$; $|p| \leqq M$ and using the inequality

$$\begin{aligned}[D(\varphi, \psi)]^2 &= [\textstyle\iint (\varphi_u\psi_u + \varphi_v\psi_v)\, dudv]^2 \\ &\leqq [\textstyle\iint (|\varphi_u\psi_u| + |\varphi_v\psi_v|)\, dudv]^2 \\ &\leqq D(\varphi)\, D(\psi)\end{aligned}$$

we find from

$$D'(p\tau) = \iint_{r\leqq\sqrt{\epsilon}} \tau^2(p_u^2 + p_v^2)\,du\,dv + \iint_{r\leqq\sqrt{\epsilon}} p^2(\tau_u^2 + \tau_v^2)\,du\,dv$$
$$+ 2\iint_{r\leqq\sqrt{\epsilon}} p\tau(p_u\tau_u + p_v\tau_v)\,du\,dv$$

---

[13] If P is the point at infinity we choose with polar coordinates around the origin

$$\tau(r) = 0 \text{ for } r \geqq 1/\epsilon;\ \tau(r) = 1 \text{ for } r < 1/\sqrt{\epsilon}$$

$$\tau(r) = -2 \log (r\epsilon)/\log \epsilon \text{ for } 1/\epsilon \geqq r \geqq 1/\sqrt{\epsilon}$$

that

$$D'(p\tau) \leqq D'(p) + M^2D'(\tau) + 2M[D'(p)\,D'(\tau)]^{\frac{1}{2}}$$

but for $r \leqq \sqrt{\epsilon}$ we have $|\,p_u\,| \leqq M/h$; $|\,p_v\,| \leqq M/h$ where h is an upper bound for the distance between the circle C_ϵ and the boundary circles C_ν.[14]

Hence it follows that

$$D'(p\tau) \leqq \frac{M^2}{h^2}\pi\epsilon - \frac{M^2 4\pi}{\log \epsilon} + \frac{M^2 4\pi}{h}[\epsilon/\log(1/\epsilon)]^{\frac{1}{2}} = \sigma(\epsilon)$$

and

$$D(q\epsilon) \leqq D(p) + \sigma(\epsilon) = d + \sigma(\epsilon)$$

as was stated.

The same consideration also proves the lemma for the case of the boundary value problem for vectors.

A statement corresponding to Lemma 2, only pointing in the opposite direction (and also not used for the case $k = 1$ of Plateau's Problem in §3), is the following

LEMMA 3: *With the notation of Lemma 2 we consider for the domain $B - K_\epsilon = B_\epsilon$, where K_ϵ is the circular domain $r \leqq \epsilon$ bounded by $C_\epsilon = C^*$, the lower limit $d_o(B_\epsilon)$ of $D(q)$ if q has the same properties as in Lemma 1 and Lemma 2, but if nothing about q is prescribed on the boundary C_ϵ.*

Then there exists a quantity $\sigma(\epsilon)$ depending only on ϵ with $\sigma(\epsilon) \to 0$, as $\epsilon \to 0$ such that

$$d_o(B_\epsilon) \geqq d(B) - \sigma(\epsilon)$$

for all boundary values of q equally bounded by $|\,q\,| \leqq M$.

In other words: *Leaving out from the domain B the circular domain K_ϵ and permitting to the functions q arbitrary values at the new boundary C_ϵ of $B_\epsilon = B - K_\epsilon$ does not improve the lower limit noticeably, if ϵ is small enough.*

PROOF: In the notation of Lemma 1 and Lemma 1a we have to show

$$\min D_{B_\epsilon}(q) = D_{B_\epsilon}(p_o) = d_o(B_\epsilon) \geqq d(B) - \sigma(\epsilon) = D_B(p) - \sigma(\epsilon)$$

with

$$\sigma(\epsilon) \to 0 \quad \text{for} \quad \epsilon \to 0,$$

where p_o is the potential function regular in B_ϵ having on the circles C_ν the prescribed boundary values and on $C_\epsilon = C^*$ vanishing normal derivatives.

We put $p + v = p_o$; p is the potential function regular in B, having the prescribed boundary values and $D(p) = d$. On C_ν we have $v = 0$. Further

$$D_{B_\epsilon}(p_0) = D_{B_\epsilon}(p) + D_{B_\epsilon}(v) - 2\int_{C_\epsilon} v\frac{\partial p}{\partial r}\,ds$$

---

[14] Here and later we are making use of the fact: If the absolute value of a potential function is less than M on a circle of radius h in which the function is regular, then the derivatives in the center have an absolute value less than M/h.

where $\partial/\partial r$ means the normal derivative and s the arc length on C_ϵ. The potential function p being regular throughout B, we have everywhere $|p| \leqq M$ and therefore on C_ϵ:

$$\left|\frac{\partial p}{\partial r}\right| < \frac{M}{h},$$

where h again is an upper bound for the distance of C_ϵ to C_ν.[15] Further $|v| < 2M$ and

$$D_{B_\epsilon}(p) = d - D_{K_\epsilon}(p) \geqq d - 2\pi\epsilon^2 M^2/h^2$$

because in K_ϵ we have

$$p_u^2 + p_v^2 < 2M^2/h^2.$$

Therefore

$$D_{B_\epsilon}(p_o) > d - \sigma(\epsilon)$$

with $\sigma(\epsilon) = 2\pi\epsilon^2 M^2/h^2 + 4\pi\epsilon M^2/h$ which proves our lemma.

Obviously the same lemma holds for potential vectors instead of functions.

We further shall need the following very elementary remark:

LEMMA 4. *Let $p(\lambda)$ be a non-negative piecewise continuous function in the interval $0 \leqq a \leqq \lambda \leqq b$ with*

$$\int_a^b p(\lambda)d\lambda \leqq M.$$

Then there exists in every subinterval $\alpha \leqq \lambda \leqq t\alpha$ a value λ_0 for which $p(\lambda_0) \leqq \epsilon/\lambda_0$ with $\epsilon = M/\log t$, so that ϵ can be chosen arbitrarily small, not depending on $p(\lambda)$ but only on the bound M, if t becomes sufficiently large.

In particular we may choose, in applying this lemma, $\alpha = \delta$; $t\alpha = \sqrt{\delta}$, $t = 1/\sqrt{\delta}$ and make δ sufficiently small to obtain a prescribed small ϵ, provided that a is small enough.

PROOF. If in $\alpha \leqq \lambda \leqq t\alpha$ we had $p(\lambda) > \epsilon/\lambda$ it would follow that

$$\int_\alpha^{t\alpha} p(\lambda)\,d\lambda > \epsilon \log t = M.$$

A consequence of basic importance for our treatment of Plateau's Problem is the following

LEMMA 5. *Suppose that in a domain B of the u, v-plane bounded by k circles $C_1, \cdots, C_k$ the vector $\mathfrak{x}(u, v)$ is continuous in the domain and on its boundaries C_ν, has piecewise continuous first derivatives in B, has the bound M for the Dirichlet Integral: $D(\mathfrak{x}) \leqq M$ and maps the circles $C_1, \cdots, C_k$ in a continuous way respectively on k prescribed Jordan curves $\Gamma_1, \cdots, \Gamma_k$ in the m-dimensional $\mathfrak{x}$-space.*

[15] See footnote 14.

Let O be a point in B or outside B, C_r the part of the circle with radius r around O lying in B. Then there exists for every sufficiently small δ a value r_0 with $\delta \leqq r_0 \leqq \sqrt{\delta}$ so that on every connected arc of C_{r_0} the oscillation of the vectors $\mathfrak{x}$ does not exceed the quantity

$$\epsilon(\delta) = [4\pi M/|\log \delta|]^{\frac{1}{2}}$$

(In this the oscillation of $\mathfrak{x}$ is counted as zero if there exists no arc of C_r in B.)

In particular, for O on C and δ so small that C_r consists of a single arc, there exist two points A_1, A_2 on C at the same distance r_0 from O with $|\mathfrak{x}(A_1) - \mathfrak{x}(A_2)| \leqq \epsilon(\delta)$.

We note that ϵ only depends on the bound M and δ and that therefore our lemma applies uniformly to all vectors $\mathfrak{x}$ with $D(\mathfrak{x}) \leqq M$ and all domains B.

It may also be noted, that *exactly the same lemma* and the same proof *hold, if B* is *any domain* in the *u, v*-plane *bounded by Jordan curves,* not necessarily circles.

Proof. Because of the supposed continuity of $\mathfrak{x}$ in B and on the boundaries it is sufficient to prove our lemma not for B but for a subdomain B' bounded by circles concentric to the circles C_ν and arbitrarily near to them. Then for B' the domain integral $D(\mathfrak{x})$ can be written as a double integral. By again using the notation B, C instead of B', C' we consider with polar coordinates $r_1\vartheta$ around O and with $s = r\vartheta$,

$$p(r) = \int_{C_r} \mathfrak{x}_s^2\, ds$$

and define $p = 0$, if the circle of the radius r has no arc in our domain. Now we certainly have

$$\int_a^b p(r)\, dr \leqq D(\mathfrak{x}) \leqq M$$

where a and b are arbitrary limits. Taking $a = \delta$ and $b = \sqrt{\delta}$ we find for two arbitrary points P and Q on C_r by means of Schwarz's inequality

$$|\mathfrak{x}(P) - \mathfrak{x}(Q)|^2 \leqq \left| \int_{C_r} \mathfrak{x}_s\, ds \right|^2 \leqq 2\pi r \int_{C_r} \mathfrak{x}_s^2\, ds = 2\pi r p(r).$$

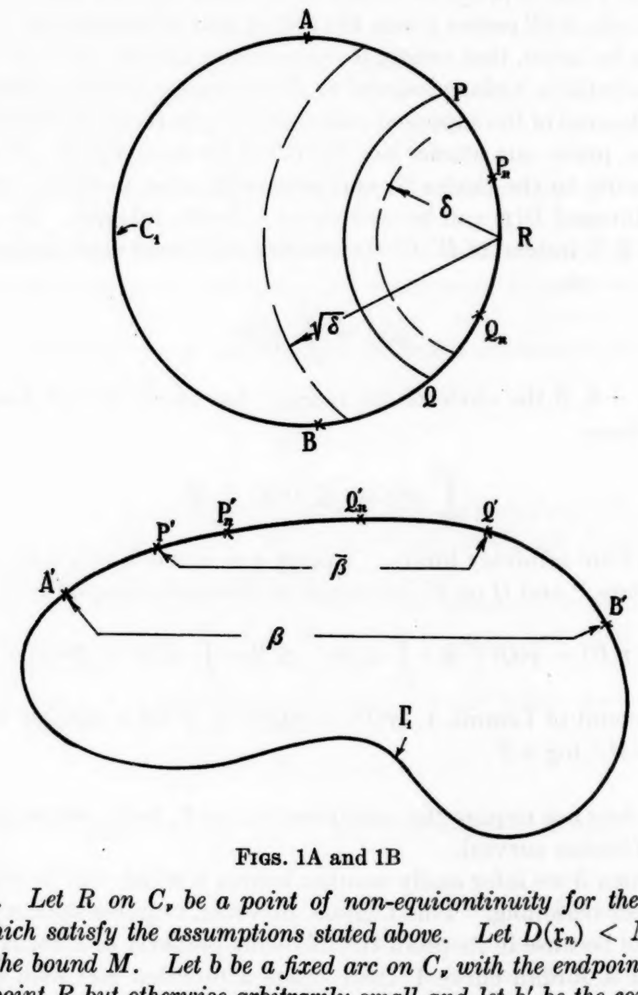
Hence on account of Lemma 4, $(\mathfrak{x}(P) - \mathfrak{x}(Q))^2 \leqq \epsilon^2$ for a suitable value $r = r_0$ with $\epsilon = [4\pi M/|\log \delta|]^{\frac{1}{2}}$.

Lemma 5 does not require the continous curves Γ_ν to be curves without multiple points (Jordan curves).

From Lemma 5 we infer easily another lemma 6 which will be very essential for our further reasoning. This Lemma, however, requires the curves Γ_ν to be Jordan curves because in its proof the following property of a Jordan curve will be used: For a Jordan curve Γ there exists a function $\eta(\epsilon)$ with $\eta(\epsilon) \to 0$ for $\epsilon \to 0$, so that any two points on Γ with the distance less than ϵ are endpoints of an arc of Γ with the diameter not exceeding $\eta(\epsilon)$.

It refers to a sequence of vectors $\mathfrak{x}_1, \mathfrak{x}_2, \cdots, \mathfrak{x}_n, \cdots$ satisfying the assumptions of lemma 5 and mapping C_ν on a Jordan curve Γ_ν in a monotone contin-

uous way. Such a sequence is called *equicontinuous* on the boundary circle C_ν of B if for any pair of points P, Q on C_ν with distance less than δ we have $|\,\mathfrak{x}_n(P) - \mathfrak{x}_n(Q)\,| < \epsilon(\delta)$ where $\epsilon(\delta) \to 0$ for $\delta \to 0$ and $\epsilon(\delta)$ independent of n. If the sequence is not equicontinuous on C_ν there exists at least one point R of *non-equicontinuity* on C_ν; that is, there exists on C_ν a sequence of points P_n and of points Q_n both converging to a point R and a positive constant α, such that $|\,\mathfrak{x}_n(P_n) - \mathfrak{x}_n(Q_n)\,| > \alpha$ for infinitely many n. By choosing a suitable subsequence and denoting it again with $\mathfrak{x}_n$ we may assume that this inequality holds for all n. Now we state



FIGS. 1A and 1B

LEMMA 6. *Let R on C_ν be a point of non-equicontinuity for the sequence of vectors $\mathfrak{x}_n$ which satisfy the assumptions stated above. Let $D(\mathfrak{x}_n) < M$ be equally bounded by the bound M. Let b be a fixed arc on C_ν with the endpoints A, B containing the point R but otherwise arbitrarily small and let b' be the complementary arc of C_ν. Then at least for a subsequence $\mathfrak{x}_n$ the image β of b defined by $\mathfrak{x}_n$ will cover all Γ_ν, except for an arc β' whose diameter tends to zero as n increases.*

In other words: *The mapping of C_ν on Γ_ν by $\mathfrak{x}_n$ tends to a degeneration in such a*

way that any small neighborhood of a point R is mapped on nearly the whole closed curve Γ_ν.

PROOF. By Lemma 5 there are for every n on C_ν, e. g. on C_1—for δ as small as we please—two points P, Q (depending on n) at the same distance r_0 from R with $\delta \leqq r_0 \leqq \sqrt{\delta}$ such that

$$|\mathfrak{x}_n(P) - \mathfrak{x}_n(Q)| < \epsilon(\delta) = [2\pi M/|\log \delta|]^{\frac{1}{2}}.$$

The images P' and Q' of P and Q on Γ_1 therefore divide Γ_1 into two arcs $\bar{\beta}$ and $\bar{\beta}'$ one of which has a diameter less than $\eta(\epsilon)$ and therefore less than any prescribed quantity as soon as δ is small enough. Now after choosing δ we choose n so large that the distance of the points P_n and Q_n described above, from R is less than δ. Since $\bar{\beta}$, as the image of the arc $P\,R\,Q$, must for fixed δ and sufficiently large n be contained in the image β of b and since the diameter of $\bar{\beta}$ exceeds α we have that the diameter of the image $\bar{\beta}'$ and all the more the diameter of the image β' of b' is less than η and tends to zero with δ; q.e.d.

COROLLARY. *The same lemma and the identical proof hold if the vector $\mathfrak{x}_n$ does not map C_ν on a fixed curve Γ_ν, but on a continuous curve $\Gamma_\nu^{(n)}$ which tends to the Jordan curve Γ_ν with increasing n.*[16] This convergence must be such that if two points A_n, B_n of $\Gamma_\nu^{(n)}$ tend to two points A, B of Γ_ν then the points in the arc A_nB_n on $\Gamma_\nu^{(n)}$ must also tend to the points of the corresponding arc AB on Γ_ν (*"strong" convergence*); but the curves $\Gamma_\nu^{(n)}$ need not necessarily be simple curves without multiple points. They may form loops which, of course, disappear in the limit.

Further it may be remarked that the same lemma holds if the vectors $\mathfrak{x}_n$ do not belong to a fixed domain B, but to a circular domain $B^{(n)}$ which converges to a circular domain B.

3. Plateau's Problem for One Contour ($k = 1$)

1. **The variational problem.** We start with the following

PROBLEM I. *Let Γ be a prescribed Jordan curve in the m-dimensional $x_1, \cdots, x_m$-space (or $\mathfrak{x}$-space); let a surface S bounded by Γ be given in parametric vector-representation by a vector $\mathfrak{x}(u, v)$ in the unit circle B of the u, v-plane with the boundary C, such that $\mathfrak{x}$ is continuous in $B + C$, maps C in a continuous and monotonic way on Γ, and has piecewise*[17] *continuous first derivatives in B. We further suppose that under this condition the Dirichlet Integral*

$$D(\mathfrak{x}) = \tfrac{1}{2}\iint_B (\mathfrak{x}_u^2 + \mathfrak{x}_v^2)\,du\,dv$$

[16] That is, each point of $\Gamma^{(n)}$ has an arbitrarily small distance from Γ if n is sufficiently large.

[17] See definition in footnote 12.

is capable of finite values.[18] *Then we ask for a vector* $\mathfrak{x}$ *for which* $D(\mathfrak{x}) = d$ *attains the lower bound* d.

We observe: *The Dirichlet Integral* $D(\mathfrak{x})$ *is invariant under conformal mapping.* Hence all the corresponding problems for other domains B' conformally equivalent to B are equivalent. E.g. we may, by a linear transformation, transform the unit circle into itself so that three given points on C are coordinated to three fixed points on Γ. Accordingly we shall assume in the following that this *three point condition* is satisfied for the vectors $\mathfrak{x}$ under consideration.

2. **Solution of Problem I.** Without knowing a priori whether a solution of Problem I exists we are sure of the existence of a minimizing sequence $\mathfrak{x}_1$, $\mathfrak{x}_2$, $\cdots$ of admissible vectors for which $D(\mathfrak{x}_n) \to d$ for $n \to \infty$ while always $D(\mathfrak{x}_n) \geqq d$. Therefore certainly these Dirichlet integrals are bounded. Now we state: The boundary values of the vectors $\mathfrak{x}_n$ are equicontinuous. Indeed, otherwise, after the choice of a suitable subsequence, we would have a point R of non-equicontinuity on C. But this, by Lemma 6 contradicts the three-point-condition, because this condition excludes the possibility of an arbitrarily small arc b of C being mapped on the whole curve Γ except for an arbitrarily small segment.

On account of the equicontinuity we can choose a subsequence of the $\mathfrak{x}_n$—again called $\mathfrak{x}_n$—which converges uniformly on the boundary C.

With these boundary values we solve the boundary value problem of the potential equation $\Delta\mathfrak{x} = 0$ for B and thus obtain a sequence of potential vectors having the same boundary values as the $\mathfrak{x}_n$ and having, on account of Lemma 1 in §1, Dirichlet Integrals not exceeding $D(\mathfrak{x}_n)$.

Since the new potential vectors also are admissible vectors in our Problem I they form all the more a minimizing sequence which we again may call $\mathfrak{x}_1$, $\mathfrak{x}_2$, $\cdots$. The uniform convergence of the boundary values of these vectors implies the uniform convergence in $B + C$ of these vectors to a potential vector. Thus, a limit vector $\mathfrak{x} = \lim_{n\to\infty} \mathfrak{x}_n$ is defined, satisfying the conditions of Problem I, and $\Delta\mathfrak{x} = 0$ in B. In each concentric circle the derivatives of $\mathfrak{x}_n$ converge uniformly towards the derivatives of $\mathfrak{x}$. Denoting by D_r the Dirichlet Integral for a concentric circle of radius $r < 1$ we have therefore for the admissible vector $\mathfrak{x}$

$$D_r(\mathfrak{x}) = \lim_{n\to\infty} D_r(\mathfrak{x}_n) \leqq \lim_{n\to\infty} D(\mathfrak{x}_n) = d.$$

Now letting r tend to 1, we obtain at once $D(\mathfrak{x}) \leqq d$, and, since the inequality sign $<$ would be in contradiction to the assumption that d is the lower bound, we have $D(\mathfrak{x}) = d$. That is: $\mathfrak{x}$ solves Problem I.

[18] This is certainly true if Γ consists of a finite number of arcs with continuous tangents. —Since later d will be recognized as the least possible area spanned by Γ the condition amounts to the existence of a surface of a finite area bounded by Γ.—By a simple passage to a limit, Douglas has shown that Plateau's Problem can be solved even if this condition is not satisfied.

Incidentally, by the same reasoning and referring to the corollary of Lemma 6 we conclude: *The lower limit d depends on the boundary* Γ *in a lower semi-continuous way.* That is: if the curves $\Gamma^{(\epsilon)}$ tend to Γ with $\epsilon \to 0$ in the strong sense and if $d(\epsilon)$ is the corresponding lower limit for $\Gamma^{(\epsilon)}$ then

$$d \leqq \lim_{\epsilon \to 0}. \inf. d(\epsilon). \tag{13}$$

3. **The solution $\mathfrak{x}$ defines a minimal surface.** The potential equation $\Delta \mathfrak{x} = 0$ is Euler's equation belonging to Problem I. The prescribed mapping of the boundary amounts to two boundary conditions. We shall show that the remaining degree of freedom in the boundary values leads to a "*natural boundary condition*"[19] which expresses the character of the surface S as a minimal surface.

In exploiting the minimizing character of the vector $\mathfrak{x}$ we need not observe the three-point condition, nor the potential character of $\mathfrak{x}$.[20] Therefore we may substitute in the Dirichlet Integral instead of $\mathfrak{x}$ another vector $\mathfrak{z}$ defined by means of concentric polar coordinates r, ϑ as follows: $\mathfrak{z}(r, \vartheta) = \mathfrak{x}(r, \varphi)$ with $\varphi = \vartheta + \epsilon\lambda(r, \vartheta)$. Here $\lambda(r, \vartheta)$ is an arbitrary function with continuous first and second derivatives in $B + C$, ϵ a small parameter. Certainly $\mathfrak{z}$ satisfies the condition of Problem I and therefore we have $D(\mathfrak{z}) \geqq d$. Now $D(\mathfrak{z})$ can be expressed in the following way by r, φ instead of r, ϑ as independent variables, if $\mathfrak{x} = \mathfrak{x}(r, \varphi)$ is substituted:

$$\begin{aligned}
2D(\mathfrak{z}) &= \int_0^1 \int_0^{2\pi} \left(\mathfrak{z}_r^2 + \frac{1}{r^2} \mathfrak{z}_\vartheta^2 \right) r\,dr\,d\vartheta \\
&= \int_0^1 \int_0^{2\pi} \left\{ (\mathfrak{x}_r + \epsilon\lambda_r \mathfrak{x}_\varphi)^2 + \frac{1}{r^2} (1 + \epsilon\lambda_\vartheta)^2 \mathfrak{x}_\varphi^2 \right\} \cdot \frac{r\,dr\,d\varphi}{1 + \epsilon\lambda_\vartheta} \\
&= \int_0^1 \int_0^{2\pi} \left(\mathfrak{x}_r^2 + \frac{1}{r^2} \mathfrak{x}_\varphi^2 \right) r\,dr\,d\varphi + \epsilon \int_0^1 \int_0^{2\pi} \left\{ 2\lambda_r \mathfrak{x}_r \mathfrak{x}_\varphi + \lambda_\vartheta \left(\frac{1}{r^2} \mathfrak{x}_\varphi^2 - \mathfrak{x}_r^2 \right) \right\} r\,dr\,d\varphi \\
&\qquad + \epsilon^2 R.
\end{aligned}$$

Here, the first term is equal to $2d$. The coefficient R of ϵ^2 remains bounded as is easily seen by Schwarz's inequality. Therefore we infer from the condition $D(\mathfrak{z}) \geqq d$, by passing to the limit $\epsilon \to 0$, that the coefficient of ϵ:

$$\int_0^1 \int_0^{2\pi} \left\{ 2\lambda_r \mathfrak{x}_r \mathfrak{x}_\varphi + \lambda_\vartheta \left(\frac{\mathfrak{x}_\varphi^2}{r^2} - \mathfrak{x}_r^2 \right) \right\} r\,dr\,d\varphi$$

tends to zero with ϵ. Again using Schwarz's inequality, for a boundary strip $\rho \leqq r \leqq 1$ and realizing that φ tends to ϑ for $\epsilon \to 0$ we obtain the relation

$$\lim_{\rho \to 1} \int_0^\rho \int_0^{2\pi} \left\{ 2\lambda_r r \mathfrak{x}_r \mathfrak{x}_\vartheta + \lambda_\vartheta r \left(\frac{\mathfrak{x}_\vartheta^2}{r^2} - \mathfrak{x}_r^2 \right) \right\} dr\,d\vartheta = 0$$

---

[19] For this concept cf. Courant-Hilbert, *Methoden der mathematischen Physik*, vol. I (1931), chapter IV, §5.

[20] Cf. for the following Radò, l.c.p. 96ff. (See also footnote 8.)

where the limit is uniform with respect to all functions λ, for which $|\lambda|, |\lambda_r|, |\lambda_\vartheta|$ are equally bounded. This equation, by the usual product integration of the calculus of variation and on account of the potential equation $\Delta\mathfrak{x} = 0$ is immediately transformed into

$$\lim_{r\to 1} \int_0^{2\pi} \lambda(r, \vartheta) r \mathfrak{x}_r \mathfrak{x}_\vartheta \, d\vartheta = 0. \tag{14}$$

Now, from the definition of the analytic function of $w = u + iv$

$$\varphi(w) = (\mathfrak{x}_u - i\mathfrak{x}_v)^2 = (E - G) - 2iF = \sum_\mu [f'_\mu(w)]^2$$

by introducing log w as new independent variable we find

$$w^2\varphi(w) = \sum_\mu \left(\frac{df_\mu}{d \log w}\right)^2 = \sum_\mu \left(r\frac{\partial x_\mu}{\partial r} - i\frac{\partial x_\mu}{\partial \vartheta}\right)^2.$$

Hence, if the symbol $\mathfrak{J}$ denotes: "imaginary part,"

$$-2r\mathfrak{x}_r\mathfrak{x}_\vartheta = \mathfrak{J}(w^2\varphi(w)) = p(u, v) = p(r, \vartheta)$$

is a potential function in B, and the equation (14) or

$$\int_0^{2\pi} \lambda(r, \vartheta) p(r, \vartheta) d\vartheta \to 0, \quad \text{as } r \to 1, \tag{15}$$

states that any linear combination of the values of this potential function p on a circle with the radius r tends to zero as r tends to 1; from this, since the value of the potential function p in any fixed interior point Q: (r_0, ϑ_0) is such a linear combination we conclude that p vanishes identically in B. Precisely, we choose for $\lambda(r, \vartheta)$ a function which vanishes identically in the neighborhood of Q and is, for r sufficiently near to 1, identical with the normal derivatives of Green's function for the concentric circle having Q as singularity (Poisson's kernel):

$$\lambda = \frac{r^2 - r_0^2}{2\pi} \frac{1}{r^2 - 2rr_0 \cos(\vartheta - \vartheta_0) + r_0^2};$$

then the equation (15) immediately becomes lim $p(Q) = 0$, that is $p(Q) = p(u, v) = 0$ everywhere in B.

Since the imaginary part p of the analytic function $w^2\varphi(w)$ vanishes, this function is real and constant in B: $w^2\varphi(w) = \text{const} = c$ or $\varphi(w) = c/w^2$. But $\varphi(w)$ is regular at $w = 0$ and therefore we have $c = 0$ or $\varphi(w) = 0$, which expresses the character of S as a minimal surface. Therewith Plateau's Problem for $k = 1$ is solved.

Once more it may be emphasized that the vector $\mathfrak{x}(u, v)$ defines a conformal mapping of the unit circle on the minimal surface S.

Further: *The minimum value d of the Dirichlet Integral is equal to the area of the minimal surface S.*

4. **Theorems on continuous dependence.** The following APPROXIMATION THEOREM proved by Douglas and Radò[21] is an immediate consequence of the preceding theory:

If the Jordan contours $\Gamma^{(\epsilon)}$ converge with $\epsilon \to 0$ to a Jordan contour Γ in the strong sense and if $\Gamma^{(\epsilon)}$ spans a minimal surface S_ϵ whose area d_ϵ is bounded by a quantity M not dependent on ϵ, then there exists a limiting minimal surface S bounded by Γ with an area not exceeding the lower limit of d_ϵ.

For, let $\mathfrak{x}_\epsilon$ be the minimizing vector which maps the unit circle B conformally on S_ϵ and satisfies a suitable three point condition then by No. 2 the equicontinuity of the vectors $\mathfrak{x}_\epsilon$ is established. Hence a subsequence of the $\mathfrak{x}_\epsilon$ can be chosen which converges uniformly to a potential vector $\mathfrak{x}$ mapping B on a surface S bounded by Γ. In this passage to the limit the relations $E - G = F = 0$ are preserved. The relation $D(\mathfrak{x}) \leqq \liminf d_\epsilon$ follows exactly as in No. 2.

In general we can, with respect to the lower limits d, state only lower semicontinuity, and not continuity, in their dependence on the boundary. But under suitable restrictions concerning the convergence of $\Gamma^{(\epsilon)}$ to Γ the lower limit d can be shown to be continuous by a simple direct reasoning.

CONTINUITY THEOREM. *Let*

$$x_i' = x_i + \xi_i(x_1, \cdots x_m) \qquad (i = 1, \cdots m)$$

be a transformation of the $\mathfrak{x}$-space into the $\mathfrak{x}'$-space with

$$|\xi_i| \leqq \epsilon, \qquad \left|\frac{\partial \xi_i}{\partial x_k}\right| \leqq \epsilon$$

transforming Γ into Γ' and the minimizing vector $\mathfrak{x}$ belonging to Γ into another vector $\mathfrak{x}'$ continuous in $B + C$, with piecewise continuous first derivatives in B and mapping B on a surface bounded by Γ'. Let the lower limits belonging to Γ and Γ' respectively be d and d'. Then we have $d' \leqq d(1 + \delta)$, with $\delta = (1 + m\epsilon)^2 - 1$ tending to zero with ϵ.

PROOF. On account of Schwarz's inequality we have

$$D(\mathfrak{x}' - \mathfrak{x}) = \sum_i \iint \left\{\left(\sum_k \frac{\partial \xi_i}{\partial x_k}\frac{\partial x_k}{\partial u}\right)^2 + \left(\sum_k \frac{\partial \xi_i}{\partial x_k}\frac{\partial x_k}{\partial v}\right)^2\right\} du dv \leqq m\epsilon^2 D(\mathfrak{x}).$$

Hence

$$(\sqrt{D(\mathfrak{x}')} - \sqrt{D(\mathfrak{x})})^2 \leqq m^2\epsilon^2 D(\mathfrak{x})$$

and therefore because of $D(\mathfrak{x}) = d$, $D(\mathfrak{x}') \geqq d'$ we have $d' \leqq (1 + m\,\epsilon)^2 d$.

5. **Riemann's mapping theorem.** As Douglas has observed, the result of No. 3 contains, for the special case $m = 2$, $(x_1 = x, x_2 = y)$ when Γ is a Jordan curve in the x, y-*plane, Riemann's mapping theorem: There exists a conformal mapping of the unit circle of the w-plane on the interior of a Jordan curve Γ in the x, y-plane.*

[21] For reference see Radò's report. Douglas also proved a more general theorem, where the boundedness of d_ϵ is not required.

where the limit is uniform with respect to all functions λ, for which $|\lambda|, |\lambda_r|, |\lambda_\vartheta|$ are equally bounded. This equation, by the usual product integration of the calculus of variation and on account of the potential equation $\Delta \mathfrak{x} = 0$ is immediately transformed into

$$\lim_{r\to 1} \int_0^{2\pi} \lambda(r, \vartheta) r \mathfrak{x}_r \mathfrak{x}_\vartheta \, d\vartheta = 0. \tag{14}$$

Now, from the definition of the analytic function of $w = u + iv$

$$\varphi(w) = (\mathfrak{x}_u - i\mathfrak{x}_v)^2 = (E - G) - 2iF = \sum_\mu [f'_\mu(w)]^2$$

by introducing log w as new independent variable we find

$$w^2\varphi(w) = \sum_\mu \left(\frac{df_\mu}{d \log w}\right)^2 = \sum_\mu \left(r\frac{\partial x_\mu}{\partial r} - i\frac{\partial x_\mu}{\partial \vartheta}\right)^2.$$

Hence, if the symbol $\mathfrak{J}$ denotes: "imaginary part,"

$$-2r\mathfrak{x}_r\mathfrak{x}_\vartheta = \mathfrak{J}(w^2\varphi(w)) = p(u, v) = p(r, \vartheta)$$

is a potential function in B, and the equation (14) or

$$\int_0^{2\pi} \lambda(r, \vartheta) p(r, \vartheta) d\vartheta \to 0, \quad \text{as } r \to 1, \tag{15}$$

states that any linear combination of the values of this potential function p on a circle with the radius r tends to zero as r tends to 1; from this, since the value of the potential function p in any fixed interior point Q: (r_0, ϑ_0) is such a linear combination we conclude that p vanishes identically in B. Precisely, we choose for $\lambda(r, \vartheta)$ a function which vanishes identically in the neighborhood of Q and is, for r sufficiently near to 1, identical with the normal derivatives of Green's function for the concentric circle having Q as singularity (Poisson's kernel):

$$\lambda = \frac{r^2 - r_0^2}{2\pi} \frac{1}{r^2 - 2rr_0 \cos(\vartheta - \vartheta_0) + r_0^2};$$

then the equation (15) immediately becomes lim $p(Q) = 0$, that is $p(Q) = p(u, v) = 0$ everywhere in B.

Since the imaginary part p of the analytic function $w^2\varphi(w)$ vanishes, this function is real and constant in B: $w^2\varphi(w) = \text{const} = c$ or $\varphi(w) = c/w^2$. But $\varphi(w)$ is regular at $w = 0$ and therefore we have $c = 0$ or $\varphi(w) = 0$, which expresses the character of S as a minimal surface. Therewith Plateau's Problem for $k = 1$ is solved.

Once more it may be emphasized that the vector $\mathfrak{x}(u, v)$ defines a conformal mapping of the unit circle on the minimal surface S.

Further: *The minimum value d of the Dirichlet Integral is equal to the area of the minimal surface S.*

4. **Theorems on continuous dependence.** The following APPROXIMATION THEOREM proved by Douglas and Radò[21] is an immediate consequence of the preceding theory:

If the Jordan contours $\Gamma^{(\epsilon)}$ converge with $\epsilon \to 0$ to a Jordan contour Γ in the strong sense and if $\Gamma^{(\epsilon)}$ spans a minimal surface S_ϵ whose area d_ϵ is bounded by a quantity M not dependent on ϵ, then there exists a limiting minimal surface S bounded by Γ with an area not exceeding the lower limit of d_ϵ.

For, let $\mathfrak{x}_\epsilon$ be the minimizing vector which maps the unit circle B conformally on S_ϵ and satisfies a suitable three point condition then by No. 2 the equicontinuity of the vectors $\mathfrak{x}_\epsilon$ is established. Hence a subsequence of the $\mathfrak{x}_\epsilon$ can be chosen which converges uniformly to a potential vector $\mathfrak{x}$ mapping B on a surface S bounded by Γ. In this passage to the limit the relations $E - G = F = 0$ are preserved. The relation $D(\mathfrak{x}) \leqq \liminf d_\epsilon$ follows exactly as in No. 2.

In general we can, with respect to the lower limits d, state only lower semicontinuity, and not continuity, in their dependence on the boundary. But under suitable restrictions concerning the convergence of $\Gamma^{(\epsilon)}$ to Γ the lower limit d can be shown to be continuous by a simple direct reasoning.

CONTINUITY THEOREM. *Let*

$$x_i' = x_i + \xi_i(x_1, \cdots x_m) \qquad (i = 1, \cdots m)$$

be a transformation of the $\mathfrak{x}$-space into the $\mathfrak{x}'$-space with

$$|\xi_i| \leqq \epsilon, \qquad \left|\frac{\partial \xi_i}{\partial x_k}\right| \leqq \epsilon$$

transforming Γ into Γ' and the minimizing vector $\mathfrak{x}$ belonging to Γ into another vector $\mathfrak{x}'$ continuous in $B + C$, with piecewise continuous first derivatives in B and mapping B on a surface bounded by Γ'. Let the lower limits belonging to Γ and Γ' respectively be d and d'. Then we have $d' \leqq d(1 + \delta)$, with $\delta = (1 + m\epsilon)^2 - 1$ tending to zero with ϵ.

PROOF. On account of Schwarz's inequality we have

$$D(\mathfrak{x}' - \mathfrak{x}) = \sum_i \iint \left\{ \left(\sum_k \frac{\partial \xi_i}{\partial x_k} \frac{\partial x_k}{\partial u} \right)^2 + \left(\sum_k \frac{\partial \xi_i}{\partial x_k} \frac{\partial x_k}{\partial v} \right)^2 \right\} du dv \leqq m\epsilon^2 D(\mathfrak{x}).$$

Hence

$$(\sqrt{D(\mathfrak{x}')} - \sqrt{D(\mathfrak{x})})^2 \leqq m^2 \epsilon^2 D(\mathfrak{x})$$

and therefore because of $D(\mathfrak{x}) = d$, $D(\mathfrak{x}') \geqq d'$ we have $d' \leqq (1 + m\,\epsilon)^2 d$.

5. **Riemann's mapping theorem.** As Douglas has observed, the result of No. 3 contains, for the special case $m = 2$, $(x_1 = x, x_2 = y)$ when Γ is a Jordan curve in the x, y-plane, *Riemann's mapping theorem: There exists a conformal mapping of the unit circle of the w-plane on the interior of a Jordan curve Γ in the x, y-plane.*

---

[21] For reference see Radò's report. Douglas also proved a more general theorem, where the boundedness of d_ϵ is not required.

The following remark may be added: In the case $m = 2$ the integral

$$A = \iint_B (x_u y_v - x_v y_u)\, dudv$$

represents for all admissible vectors $\mathfrak{x}$ in the Problem I the fixed area A included in the Jordan curve Γ. Therefore instead of attempting to minimize the Dirichlet Integral we could just as well start with the variational problem $D(\mathfrak{x}) - A = \text{min.}$ or

$$\iint_B \{(x_u - y_v)^2 + (x_v + y_u)^2\}\, dudv = \text{min.}$$

This is exactly the famous variational problem which Riemann considered in his doctor's thesis. Of course, the minimum value here is zero, and accordingly the solution satisfies the Cauchy-Riemann equations.

Riemann's mapping theorem for the interior of Jordan curves is obtained by this method merely on the basis of the knowledge of Poisson's integral. Moreover, our method establishes directly: *The conformal mapping of the interior B of the unit circle on the interior G of a Jordan curve Γ implies a continuous one-to-one mapping of the boundaries C and Γ* (See also No. 6). Indeed, that C is mapped continuously on Γ was shown above as a consequence of lemma 5. Conversely, to each point on Γ corresponds only one point on C. This also follows from lemma 5. For, if the inverse conformal mapping of G on B is given by the functions $u(x, y)$, $v(x, y)$ we have

$$\iint_G (u_x^2 + u_y^2 + v_x^2 + v_y^2)\, dxdy = 2\pi.$$

Now we consider a point Q on Γ and small circular arcs C_r with the radius r around Q which together with an arc of Γ bound simply connected subdomains G_r of G having Q as boundary point. On account of lemma 5 we can make the diameter of the image of C_r in B arbitrarily small for sufficiently small values of r. Hence the diameter of the image B_r of G_r must tend to zero together with r. Consequently the nested domains B_r define a single limiting point P on C which corresponds to Q.[22]

6. The solution of Plateau's problem furnishes a one-to-one mapping of the boundaries. For the general case of Plateaus problem with $m > 2$ the one-to-one correspondence of the boundaries can be proved as follows:

Since according to our construction the mapping of C on Γ is continuous and

---

[22] For the reader who is interested in conformal mapping as such, it may be stated that the same reasoning subsists for the mapping of the unit circle on a general domain not necessarily bounded by a Jordan curve. Then our circular arcs C_r define nested domains G_r in G whose limiting points form a "*prime end,*" and it is such a prime end to which a single point P on C corresponds. Our reasoning establishes the one-to-one correspondence between the prime-ends of G and the points of C.

eo ipso monotonic we only have to prove that a whole arc b on C never corresponds to the same point, e.g. $\mathfrak{x} = 0$, on Γ.

If b corresponds to the point $\mathfrak{x} = 0$ we could, by an elementary transformation, map B conformally on a half unit circle B' so that b corresponds to the diameter PQ. The minimizing vector is transformed into a vector $\mathfrak{x}'$ in B' for which, because of the invariance of the Dirichlet Integral, $D_{B'}(\mathfrak{x}') = d$.

Now we complete B' to the whole unit circle by affixing the complementary half circle B'' and define $\mathfrak{x}' = 0$ in B''. Then for the unit circle $B = B' + B''$ we have $D(\mathfrak{x}') = d$, and $\mathfrak{x}$ is an admissible vector in our variational problem I. Since $\mathfrak{x}'$ is certainly not a regular potential vector in the whole circle B we have another admissible potential vector $\mathfrak{z}$ with the same boundary values and with $D(\mathfrak{z}) < d$ (Lemma 1), which contradicts the minimum-character of d. Therefore no such arc b can occur, and our statement is proved.[23]

4. Plateau's Problem for k Contours. Preparations

1. **Heuristic considerations.** If the boundary Γ of the required minimal surface S consists of k oriented Jordan curves $\Gamma_1, \cdots, \Gamma_k$ and S is to be mapped on a domain B in the u, v-plane with the boundary C consisting of k correspondingly oriented curves $C_1, \cdots, C_k$, then we have no longer complete freedom in the choice of this domain. For, two domains, in particular two circular domains, are conformally equivalent, if and only if certain conformal invariants, the moduli, coincide. Without basing our theory on the knowledge of this fact we therefore will in our variational problem leave sufficient leeway for the choice of our circular domain B.

In the case $k > 1$, we must expect that the problem itself requires certain conditions for its solvability. For, if the curves are too far apart or otherwise misplaced, the corresponding least area problem may have a solution consisting of several separated surfaces each bounded by only a part of the system of curves Γ. In this case for the lower limit δ of the area in consideration we have $\delta = \delta' + \delta''$ where δ' and δ'' are the corresponding lower limits for the system Γ' consisting of one part of the Γ_ν and for the complementary system Γ''.

It is easily seen that under all circumstances the inequality $\delta \leqq \delta' + \delta''$ holds. Indeed, a system of two surfaces S' and S'' bounded by Γ', Γ'' resp. always can be modified into a surface S with Γ as boundary by joining S' and S'' by a tiny pipelike connection with arbitrarily small area; therefore the lower limit corresponding to Γ cannot exceed the sum $\delta' + \delta''$.

In the case $\delta = \delta' + \delta''$ a minimizing sequence $s_1, s_2, \cdots$ of surfaces bounded

---

[23] This proof does not make use of the character of our solution as a minimal surface. Taking advantage of the result $E - G = F = 0$ Douglas gives another proof in the following way: Since the potential vector $\mathfrak{x}$ is constant on b it can be analytically extended, by the principle of reflection, beyond b, and therefore is regular on b. If s denotes the arc length on b we have $\mathfrak{x}_r^2 - \mathfrak{x}_s^2 = 0$. But since $\mathfrak{x}_s = 0$ on b, the consequence is $\mathfrak{x}_r = 0$ on b. This implies that on b all the first derivatives of the potential vector $\mathfrak{x}$ vanish which leads to the absurd consequence that $\mathfrak{x}$ is constant throughout its domain of existence.

by Γ and with areas tending to δ must be expected to have the tendency to degenerate into at least two separate parts in the following way: On S_n there exists a closed curve T_n with its diameter tending to zero, so that T_n separates S_n into two parts, one bounded by Γ', the other by Γ''.

These heuristic considerations for the area do not apply immediately to our Dirichlet Integral. We therefore will give in No. 3 and in the following section a precise formulation and a proof of corresponding statements for the Dirichlet Integral.

2. The variational problem. We consider in the u, v-plane a circular domain B with boundary C consisting of k circles $C_1, \cdots, C_k$. Whether this domain is enclosed in one of these circles C_1 or contains the infinite point of the plane of the complex variable $w = u + iv$ does not matter. In the domain B again $\mathfrak{x}(u, v)$ denotes a vector, continuous in $B + C$, having piecewise continuous first derivatives in B and mapping the boundary circles monotonically in the prescribed sense on the prescribed Jordan curves $\Gamma_1, \cdots, \Gamma_k$ in the m-dimensional space. The lower limit of the Dirichlet Integral for all these admissible vectors $\mathfrak{x}$ and fixed B is called $d(B)$. The corresponding lower limit, if not only $\mathfrak{x}$, but also the circular domain B, can be chosen arbitrarily, may be called d, the "*absolute minimum*." (d is therefore the lower limit of the values $d(B)$). Now we again consider

Problem I. *To find a circular domain B and in it an admissible vector $\mathfrak{x}$ for which $D(\mathfrak{x})$ attains its absolute minimum d.* We expressly suppose as in the case $k = 1$, that our boundary Γ allows finite values of the Dirichlet Integral.

If in our variational problem the domain B is fixed we shall occasionally denote it by Problem I′ or by Problem I(B).

It may again be emphasized that because of the invariance of the Dirichlet Integral under conformal mapping we can replace B by any domain obtained by a linear transformation of the w-plane. In particular we can choose one of the circles C_ν as the unit circle and establish on this unit circle a three-point condition just as for $k = 1$. Or we can require that two circles C_1 and C_2 be concentric, C_1 being the unit circle, and that one fixed point on C_1 be transformed into a fixed point on Γ_1. Or we may replace a domain B contained in a circle C_1 by another circular domain which contains the infinite point.

We shall assume B inside the unit circle C_1. Then we consider a subset C' of the set of boundary circles, and among them C_1. The corresponding curves $\Gamma_1, \cdots$ form a subset Γ' of the boundary Γ. The circles C' define a circular domain B' included in C_1. The complementary set C'' of boundary circles C_ν defines a circular domain B'' containing the point at infinity, and the corresponding curves Γ_ν form a subset Γ'' of the boundary Γ so that we have $C' + C'' = C$, $\Gamma' + \Gamma'' = \Gamma$. The original domain B is the product of B' and B'', that is the point set common to these two domains..

If in our variational problem I or I(B) we replace the domain B by B' or B''

and Γ by Γ' or Γ'' we have to consider lower limits $d(B')$, $d(B'')$ for fixed domains and the *absolute minima* d' or d'' which depend only on Γ' or Γ'' respectively.

3. Lemmas concerning the lower limits of $D(\mathfrak{x})$.

LEMMA a. *If Γ' and Γ'' are two complementary sets of boundary curves Γ, i.e. $\Gamma' + \Gamma'' = \Gamma$ and if d, d' and d'' are the corresponding absolute minima then we have $d \leqq d' + d''$.*

PROOF. We consider a circular domain B' included in the unit circle C_1 so that there exists an admissible potential vector $\mathfrak{x}'$ which maps B' on a surface bounded by Γ' and which for a prescribed small δ satisfies $D_{B'}(\mathfrak{x}') < d' + \delta$. If O is an arbitrary point in B', e.g. the origin, we cut out from B' a concentric circular disc K_ϵ bounded by the small circle C^* with the radius ϵ around O. For the remaining domain $B'_\epsilon = B' - K_\epsilon$ we consider the new variational problem obtained by imposing the new boundary-condition $\mathfrak{x} = 0$ on C^*. By Lemma 2 we have for the corresponding lower limit $d'(B'_\epsilon)$ the inequality

$$d'(B'_\epsilon) < D(\mathfrak{x}') + \sigma(\epsilon) < d' + \delta + \sigma(\epsilon)$$

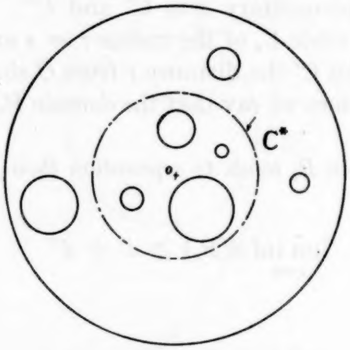


FIG. 2

where $\sigma(\epsilon) \to 0$ for $\epsilon \to 0$. Accordingly we have in B'_ϵ an admissible vector $\mathfrak{z}'$ for which

$$D_{B'_\epsilon}(\mathfrak{z}') \leqq d' + \delta + \sigma(\epsilon).$$

In the same way we consider a domain B'' containing the point at infinity; from B'' we cut out the exterior of a sufficiently large circle with the radius $1/\epsilon$ around O; in the remaining domain B''_ϵ we have a vector $\mathfrak{z}''$ which vanishes at the circumference of the large circle, which satisfies on the other boundary circles and in B''_ϵ the conditions for admission in our variational problem and the condition

$$D_{B''_\epsilon}(\mathfrak{z}'') \leqq d'' + \delta + \sigma(\epsilon).$$

Now by a similarity transformation with the factor ϵ^2 we contract B''_ϵ into a domain included in the circle C^* with the radius ϵ around O and again call the new domain B''_ϵ, and the corresponding vector $\mathfrak{z}''$. The value of the Dirichlet

integral $D_{B_\epsilon''}(\mathfrak{z}'')$ is not changed by this transformation. For the domain $B = B_\epsilon' + B_\epsilon''$ we have certainly, if $\mathfrak{z}$ denotes $\mathfrak{z}'$, $\mathfrak{z}''$ respectively,

$$D_B(\mathfrak{z}) < d' + d'' + 2\delta + 2\sigma(\epsilon).$$

If we impose on the vectors in the variational problem defining $d(B)$ the additional condition $\mathfrak{x} = 0$ on C^* and therefore narrow the range of competition, we obtain a lower limit $d(B, \epsilon)$ which certainly is not smaller than $d(B)$. But our condition $\mathfrak{x} = 0$ on C^* obviously establishes two independent variational problems for B_ϵ' and B_ϵ'' and $\mathfrak{x} = \mathfrak{z}$ is admissible. Therefore we have

$$d \leqq d(B) \leqq d(B, \epsilon) \leqq D_B(\mathfrak{z}) \leqq d' + d'' + 2\delta + 2\sigma(\epsilon)$$

and since δ as well as ϵ could be chosen arbitrarily small while d does not depend on δ and ϵ we have $d \leqq d' + d''$ as stated in the Lemma.

The considerations of the preceding Lemma are supplemented by another Lemma pointing in the opposite direction and useful later for $k > 2$ in excluding the degeneration of our parameter-domain. We consider again a finite circular domain $B = B_\epsilon$ with the following property: the set of boundary circles C is divided into two complementary sets C' and C''. One of these sets C'' consists of circles inside a circle C_ϵ of the radius $r = \epsilon$ around a fixed point O of B. For all the other circles C' the distance r from O shall remain above a fixed quantity h. If ϵ tends to zero we say that the *domain B_ϵ tends to separation.*

Lemma b. *If the domain B_ϵ tends to separation then for the lower limit of the values $d(B_\epsilon)$ the inequality*

$$\liminf_{\epsilon \to 0} d(B_\epsilon) \geqq d' + d''$$

holds. Precisely

$$d(B_\epsilon) \geqq d' + d'' - \sigma(\epsilon)$$

where $\sigma(\epsilon)$ tends to zero uniformly for a fixed bound h and for a fixed bound M for $|\mathfrak{x}|$ on Γ. In the particular case $k = 2$ the more precise limiting equation

$$\lim_{\epsilon \to 0} d(B_\epsilon) = d' + d''$$

is true.

Proof. Let B' be the finite domain bounded by C'. Removing the interior K_η of the circle $C_\eta = C^*$ with the radius $r = \eta = \sqrt{\epsilon}$ around O from B' we obtain the domain $B' - K_\eta = B_\eta'$. Now we consider the lower limit $d_0(B_\eta')$ of the Dirichlet Integral $D_{B_\eta'}(\mathfrak{x})$ if the vector $\mathfrak{x}$ in B_η' maps C' on Γ', but does not have prescribed boundary values on C^*. Lemma 3 of §2 states that

$$d_0(B_\eta') \geqq d(B') - \sigma'(\eta) \geqq d' - \sigma'(\eta).$$

where $\sigma'(\eta)$ tends to zero with η uniformly for a fixed bound h and for a fixed bound M for $|\mathfrak{x}|$ on Γ.

In the same way we obtain with corresponding notations

$$d_0(B''_\eta) \geqq d(B'') - \sigma''(\eta) \geqq d'' - \sigma''(\eta)$$

where the index η refers to a domain bounded by C'' and the circle $C_\eta = C^*$ and where the boundary values on C^* of the vectors in competition are free.

Now we return to our variational problem I (B_ϵ) defining $d(B_\epsilon)$. While in the proof of Lemma a we increased the lower limit by narrowing the range of competition, we now widen the range of competition by easing the conditions for admission, and we therefore obtain a not larger lower bound in the following way:

We permit the admissible vectors $\mathfrak{x}$ to be discontinuous at the circle $C_\eta = C^*$. In this new variational problem the two domains B'_η and B''_η are entirely separate and therefore the new lower limit is equal to the sum of the corresponding lower

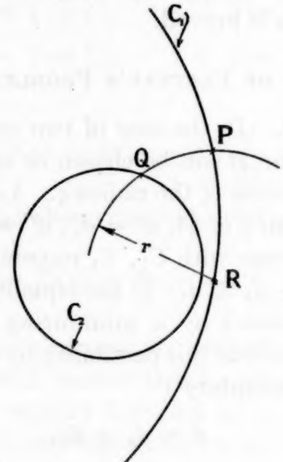


FIG. 3

limits $d_0(B'_\eta)$ and $d_0(B''_\eta)$. We therefore obtain, considering the inequalities above,

$$d(B_\epsilon) \geqq d_0(B'_\eta) + d(B''_\eta) \geqq d' + d'' - \sigma'(\eta) - \sigma''(\eta),$$

which for $\epsilon \to 0$ and $\sigma(\epsilon) = \sigma'(\eta) + \sigma''(\eta)$ contains the statement of our lemma.[23a]

LEMMA c. *If in a sequence of domains B_ϵ with $\epsilon \to 0$ two boundary circles C_1, C_2 with radii above a fixed bound come closer than ϵ, then $\lim_{\epsilon\to 0} d(B_\epsilon) \to \infty$; that is the lower limits $d(B_\epsilon)$ are not bounded.*

PROOF. Without restricting the generality we may suppose that C_1 is the unit circle and C_2 approaches, not shrinking to zero, a fixed point R on C_1.

[23a] For $k = 2$ we know from the considerations of Lemma a) that $d(B_\epsilon) \leqq d(B') + d(B'') + \sigma(\epsilon)$ with $\sigma(\epsilon)$ tending to zero. Hence in this case we really have the equality sign $\lim_{\epsilon\to 0} d(B) = d(B') + d(B'') = d' + d''$.

Then for sufficiently small δ all circles with radii r between a fixed, small, δ and $\sqrt{\delta}$ around R intersect C_1 and C_2 in two points P and Q. Since C_1 is mapped on Γ_1 and C_2 on Γ_2 we certainly have for every admissible vector $\mathfrak{x}$ in B_ϵ that $|\mathfrak{x}(P) - \mathfrak{x}(Q)| \geqq \alpha$ where α is the positive distance between Γ_1 and Γ_2. But on account of Lemma 5, §2 there exists a pair of points P, Q for which, with $D_{B_\epsilon}(\mathfrak{x}) = M_\epsilon$,

$$|\mathfrak{x}(P) - \mathfrak{x}(Q)| \leqq \left[\frac{2\pi M_\epsilon}{|\log \delta|}\right]^{\frac{1}{2}},$$

hence

$$M_\epsilon \geqq \frac{\alpha^2}{2\pi}|\log \delta|.$$

This shows that also $d(B_\epsilon) \geqq \frac{1}{2}(\alpha^2/\pi) \cdot |\log \delta|$, and since δ can be chosen arbitrarily small, our lemma is proved.

5. Solution of Plateau's Problem for $k = 2$

1. **Solution of Problem** I. In the case of two contours ($k = 2$) the solution proceeds very simply because B can be chosen as an annular ring, C_1 being the unit circle, C_2 a concentric circle of the radius q. Let d be the lower limit of the Dirichlet Integral in Problem I of §4, $d' = d_1$, $d'' = d_2$ the corresponding lower limits for the one-contour case with C_1, Γ_1 respectively C_2, Γ_2 as boundaries. By Lemma a we have $d_1 + d_2 \geqq d$. If the equality sign holds, then the lower limit is certainly approximated by a minimizing sequence of surfaces tending to degeneration. We exclude this possibility by expressly making in Problem I the assumption[24] for the boundary Γ

$$d < d_1 + d_2. \tag{16}$$

Now let $\mathfrak{x}_1, \mathfrak{x}_2, \cdots$ be a minimizing sequence of vectors; $B_1, B_2 \cdots$ a corresponding sequence of annular rings. Lemma b of §4 shows that the radii q_ν of C_2 cannot have zero as lower limit, because for such sequences of rings the lower limit $d(B_\nu)$ must be not less than $d_1 + d_2$, while according to our assumption (16), $\lim D_{B_\nu}(\mathfrak{x}_\nu) < d_1 + d_2$.

[24] This condition can be interpreted as follows. We consider on a surface S bounded by Γ_1 and Γ_2 all closed curves topologically equivalent on S to these boundary curves. Let the lower limit of their diameters be α. Then, if we impose in Problem I the condition $\alpha = 0$, the lower limit will be $d_1 + d_2$. Our inequality now requires that the additional condition $\alpha = 0$ increases the value of the lower limit d. This has been generalized by M. Shiffman (See also footnote 9) in the following way: Instead of Problem I we consider a Problem Ia by requiring that for our admissible surfaces we have $\alpha \geqq a$ where a is any given number. If then by imposing the new condition $\alpha = a$ the lower limit is actually increased, it can be shown that Problem Ia has a solution which gives a minimal surface with $\alpha > a$. This solution may be different from that obtained by Problem I.

The same principle can be applied to obtain other solutions of Plateau's problem if we replace α by another suitable continuous functional of the surface $\mathfrak{x}(u, v)$.

On account of Lemma c in §4 no subsequence of the q_ν can tend to 1 because the Dirichlet integrals $D(\mathfrak{x}_\nu)$ are bounded. Hence at least for a subsequence of the ν the domains B_ν tend to a domain B where B is an annular ring between the unit circle C_1 and a circle C_2 having a radius q with $0 < q < 1$.

The boundary values of $\mathfrak{x}_\nu$ on the boundaries C_1 and C_2 of B_ν are equicontinuous.

Proof. If this were not true e.g. on C_1 we could, on account of Lemma 6 of §2, find at least a subsequence of the $\mathfrak{x}_\nu$ for which an arbitrarily small fixed arc b on C_1 containing a fixed point R would be mapped by $\mathfrak{x}_\nu$ on an arc $\Gamma_1 - \gamma_\nu$ nearly covering Γ_1 entirely; γ_ν is a small part of Γ_1 tending to a single point, which we may assume, without loss of generality, to be the point $\mathfrak{x} = 0$. We also may assume that the minimizing sequence $\mathfrak{x}_\nu$ consists of potential vectors: $\Delta\mathfrak{x}_\nu = 0$. Now we put $\mathfrak{x}_\nu = \mathfrak{x} = \mathfrak{y} + \mathfrak{z}$ where $\mathfrak{y}$ and $\mathfrak{z}$ are also potential vectors and where

$$\mathfrak{y} = \mathfrak{x} \text{ on } C_1; \qquad \mathfrak{z} = 0 \text{ on } C_1$$
$$\mathfrak{y} = 0 \text{ on } C_2; \qquad \mathfrak{z} = \mathfrak{x} \text{ on } C_2 .$$

We have

$$D(\mathfrak{x}_\nu) = D(\mathfrak{x}) = D(\mathfrak{y}) + D(\mathfrak{z}) + 2\int\int_{B_\nu} (\mathfrak{z}_u\mathfrak{y}_u + \mathfrak{z}_v\mathfrak{y}_v)\,du\,dv.$$

Applying Green's formula and observing that $\mathfrak{z} = 0$ on C_1, $\mathfrak{y} = 0$ on C_2, we obtain

$$D(\mathfrak{x}_\nu) = D(\mathfrak{y}) + D(\mathfrak{z}) - 2\int_{C_1} \mathfrak{x}\frac{\partial\mathfrak{z}}{\partial r}\,ds \tag{17}$$

where the last integral is extended over the contour C_1, s denotes the arc length on C_1, $\partial/\partial r$ differentiation with respect to the radius r.

On this circle $|\,\partial\mathfrak{z}/\partial r\,|$ is equally bounded by a certain bound M. For the components of $\mathfrak{z}$ vanish on C_1 and hence this vector can be analytically extended as a potential vector, by the principle of reflection, beyond C_1 into a concentric ring between the radii q and $1/q$. Since the components of $\mathfrak{z}$ are bounded in this ring by $|\,\mathfrak{z}\,| \leqq A$, where A is the greatest distance of Γ from the origin, the components of the derivatives on C_1 have the bound $A/(1 - q) = M$. For the values $|\,\mathfrak{x}\,|$ on b we have just as well the bound A, while on the complementary arc $C_1 - b$ we have with arbitrarily small ϵ and for sufficiently large ν the inequality $|\,\mathfrak{x}\,| \leqq \epsilon$. Therefore, if l is the length of b,

$$\left|\,2\int_{C_1} \mathfrak{x}\frac{\partial\mathfrak{z}}{\partial r}\,ds\,\right| < \frac{2A^2 l}{1 - q} + 2\epsilon\frac{A}{1 - q}\cdot 2\pi = \sigma(\nu).$$

Since l can be chosen arbitrarily small as well as ϵ, if ν is large enough, $\sigma(\nu)$ tends to zero as ν increases. Hence we have from (17)

$$D(\mathfrak{x}_\nu) \geqq D(\mathfrak{y}) + D(\mathfrak{z}) - \sigma(\nu)$$

or since $D(\mathfrak{y}) \geqq d_1$, $D(\mathfrak{z}) \geqq d_2$,

$$D(\mathfrak{x}_\nu) \geqq d_1 + d_2 - \sigma(\nu);$$

and for $\nu \to \infty$ $d \geqq d_1 + d_2$, in contradiction to our assumption (16). This proves the equicontinuity of the boundary values of $\mathfrak{x}_\nu$ on C_1, and similarly on C_2.

On account of this fact we can choose a subsequence of the $\mathfrak{x}_\nu$ for which the boundary values converge uniformly. Therefore the corresponding potential vectors also converge uniformly in B to a potential vector $\mathfrak{x}$ which maps C on Γ and for which, with the same reasoning as in §3, we have $D(\mathfrak{x}) \leqq d$ and therefore, since d is the lower limit for admissible vectors and since $\mathfrak{x}$ is admissible, $D(\mathfrak{x}) = d$.

Thus, Problem I is solved by a potential vector $\mathfrak{x}$, or by the surface S defined by $\mathfrak{x}$.

2. The solution furnishes a minimal surface S. Since C_1, C_2 are concentric circles we can proceed as in §3 for $k = 1$. We perform a variation of the minimizing vector $\mathfrak{x}(r, \vartheta)$ (in concentric polar coordinates) again replacing $\mathfrak{x}$ by $\mathfrak{z}(r, \vartheta) = \mathfrak{x}(r, \varphi)$, with $\varphi = \vartheta + \epsilon\lambda(r, \vartheta)$. Thus we find exactly as before,

$$\lim_{\substack{r_1 \to 1 \\ r_2 \to q}} \left\{ \int_0^{2\pi} \lambda(r_1; \vartheta) r_1 \mathfrak{x}_r \mathfrak{x}_\vartheta \, d\vartheta - \int_0^{2\pi} \lambda(r_2; \vartheta) r_2 \mathfrak{x}_r \mathfrak{x}_\vartheta \, d\vartheta \right\} = 0. \tag{18}$$

As before in §3, No. 3 we conclude from this equation: The analytic function $\varphi(w) = E - G - 2iF$ is expressed in B by $\varphi(w) = c/w^2$ where c is a real constant. For: we may choose λ in the neighborhood of the two boundary circles as identical with the normal derivative of Green's function for a fixed singular point Q in B and the ring $r_2 \leqq r \leqq r_1$. Then (18) expresses the fact that the imaginary part of $w^2\varphi(w)$ vanishes in Q.[25]

We have to show that in the equation $w^2\varphi(w) = c$ the constant c vanishes. Here our former argument of §3 collapses because the point $w = 0$ does not belong to B. But we obtain the desired relation $c = 0$ as the variational condition, not yet exploited, which states that we cannot reach a smaller lower limit d by variation of the domain B, i.e. of the radius q. This variation with respect to the size of the circle C can be performed by the following method.

To get rid of the difficulty which arises from our lack of knowledge about the derivatives of $\mathfrak{x}$ at the boundaries, we replace the domain B by another concentric annular ring B' between the circles C_1': $r = r_1 < 1$ and C_2': $r = r_2 = \rho > q$; r, ϑ again are concentric polar coordinates. We show first that the minimum property of our solution with respect to the domain B implies a similar property with respect to B'. The potential vector $\mathfrak{x}(r, \vartheta)$ which solves Problem I is analytic on C_1' and C_2' so that the values $\mathfrak{x}(r_1; \vartheta) = f_1(\vartheta)$ and $\mathfrak{x}(\rho, \vartheta) = f_2(\vartheta)$ are analytic functions of the angle ϑ. The domain B consists of the domain B' plus the two annular rings R_1: $r_1 \leqq r \leqq 1$ and R_2: $q \leqq r \leqq \rho$. Now we perform a

[25] If we do not want to make use of Green's function for a ring, but only of Poisson's kernel we may choose λ in the neighborhood of one of the circles as this Poisson kernel and as zero otherwise; then our equation yields (cf. the general case discussed later in §6, No. 2) that the imaginary part of $w^2\varphi(w)$ vanishes first on C_1 and second on C_2, therefore identically in B.

variation with the annular ring B' replacing the inner circle C_2' by another circle $C_2^*: r = \sigma = t\rho$. In the new ring B^* we define a potential vector $\mathfrak{z}(r, \vartheta, \sigma) = \mathfrak{z}(u, v, \sigma)$ as the vector which has the boundary values $f_1(\vartheta)$ on C_1' and $f_2(\vartheta)$ on C_2^*. Obviously we have $\mathfrak{z}(u, v, \rho) = \mathfrak{x}(u, v)$. Now we state: The Dirichlet Integral

$$D^*(\mathfrak{z}) = \tfrac{1}{2}\int\int_{B^*} (\mathfrak{z}_u^2 + \mathfrak{z}_v^2)\, du dv = D(\sigma)$$

(where the superscript* indicates that B^* is the domain of integration) is, as a function of σ, a minimum for $\sigma = \rho$; $B^* = B'$. (In other words: *The minimum property obvious for the original ring B itself holds also for concentric rings.*)

To prove our statement we first affix to our domain B^* the original ring R_1 again; further we affix the annular ring $R_2^*: tq \leqq r \leqq t\rho = \sigma$ to the inner circle C_2^* of B^*. R_2 is transformed into R_2^* just by a dilation. Now we consider in the ring domain $G = B^* + R_1 + R_2^*$ the vector $\mathfrak{z}$ which is identical with $\mathfrak{x}$ in R_1, to

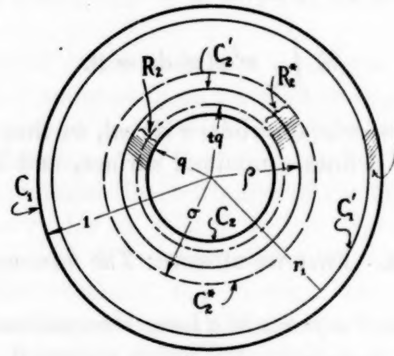


FIG. 4

$\mathfrak{z}(u, v, \sigma)$ in B^* and whose value at every point of R_2^* is identical with the value of $\mathfrak{x}$ at the corresponding point of R_2. The new domain G and in it the vector $\mathfrak{z}$ are admissible for competition in Problem I. Therefore we have

$$D_G(\mathfrak{z}) = D_{R_1}(\mathfrak{z}) + D_{R_2^*}(\mathfrak{z}) + D^*(\mathfrak{z}) \geqq d = D_{R_1}(\mathfrak{x}) + D_{R_2}(\mathfrak{x}) + D_{B'}(\mathfrak{x});$$

but according to our definition, $D_{R_1}(\mathfrak{z}) = D_{R_1}(\mathfrak{x})$ and $D_{R_2^*}(\mathfrak{z}) = D_{R_2}(\mathfrak{x})$. Therefore it follows immediately that $D^*(\mathfrak{z}) \geqq D_{B'}(\mathfrak{x})$, which is our statement.

Now since our potential vector $\mathfrak{z}$ depends analytically on the parameter σ the same is true for the Dirichlet Integral $D^*(\mathfrak{z}) = D(\sigma)$. Hence we can differentiate with respect to the parameter σ according to the rules of the elementary integral calculus. Thus we obtain the equation

$$0 = \frac{\partial}{\partial\sigma} D(\sigma)\,\Big|_{\sigma=\rho} = -\int_{r=\rho}^{2\pi} \rho(\mathfrak{x}_u^2 + \mathfrak{x}_v^2)\, d\vartheta + 2\int\int_{B'} (\mathfrak{z}_u \mathfrak{z}_{u\sigma} + \mathfrak{z}_v \mathfrak{z}_{v\sigma})\, du dv\,\Big|_{\sigma=\rho}.$$

By using Green's formula, putting $ds = \rho d\vartheta$ and observing that $\mathfrak{z}_\sigma = 0$ on C_1', we transform this into

$$0 = -\int_{r=\rho} (\mathfrak{x}_u^2 + \mathfrak{x}_v^2)\, ds - 2\int_{r=\sigma=\rho} \lambda \mathfrak{x}_r\, ds \quad \text{with} \quad \lambda = \mathfrak{z}_\sigma\big|_{\sigma=\rho}.$$

Now $\mathfrak{x}_u^2 + \mathfrak{x}_v^2 = \mathfrak{x}_r^2 + \mathfrak{x}_s^2$ and according to the definition of $\mathfrak{z}(u, v, \sigma)$

$$\lambda = \mathfrak{z}_\sigma(\rho, \vartheta, \sigma)\,\big|_{\sigma=\rho} = -\,\mathfrak{x}_r\,\big|_{r=\rho}\,.$$

Hence

$$\int_{r=\rho} (\mathfrak{x}_r^2 - \mathfrak{x}_s^2)\,ds = 0. \tag{19}$$

By virtue of $w^2\varphi(w) = r^2(\mathfrak{x}_r^2 - \mathfrak{x}_s^2) - 2ir^2\mathfrak{x}_r\mathfrak{x}_s$ our formula (19) becomes

$$\mathfrak{R}\int_{|w|=\rho} w^2\varphi(w)\,ds = 0$$

where $\mathfrak{R}$ denotes the real part.

Therefore our variation of the domain B results in the following fundamental condition:

For every concentric circle $r = \rho$ we have

$$\mathfrak{R}\int_{r=\rho} w^2\varphi(w)\,ds = 0. \tag{20}$$

Substituting in (20) $c = w^2\varphi(w)$, where c is real, we find immediately $c = 0$ or $\varphi(w) = 0$. Therefore $\mathfrak{x}$ defines a minimal surface, and Plateau's Problem for $k = 2$ is solved.

3. **Additional remarks.** First we observe: *The minimum d is the area of the minimal surface S.*

Further: *The minimum d depends in a lower semicontinuous way on the boundaries Γ.* Here again the boundaries $\Gamma^{(\epsilon)}$ which approach Γ in the strong sense as ϵ tends to zero need not consist of simple curves. To prove the statement we may assume that the minima d_ϵ belonging to the boundaries $\Gamma^{(\epsilon)}$ are bounded in ϵ. We consider a subsequence for ϵ such that $d_\epsilon \to \underline{d}$ where the bar indicates the lower limit with respect to ϵ. We have to prove $\underline{d} \geqq d$.

Now we have two alternative possibilities

1) $d_\epsilon > d'_\epsilon + d''_\epsilon - \lambda(\epsilon)$ with $\lambda \to 0$ for $\epsilon \to 0$

2) There exists a positive fixed α and a subsequence of ϵ tending to zero such that for this subsequence $d_\epsilon < d'_\epsilon + d''_\epsilon - \alpha$.

In the first case we have immediately for the lower limits the relation $\underline{d} \geqq \underline{d}' + \underline{d}''$ since for $k = 1$ already $\underline{d}' \geqq d'$, $\underline{d}'' \geqq d''$ is known from §3, No. 2 and since $d' + d'' > d$ on account of Lemma a, the relation $\underline{d} > d$ is established.

In the second case we consider for our sequence ϵ the minimizing potential vectors $\mathfrak{x}_\epsilon$ belonging to the variational problem I for $\Gamma^{(\epsilon)}$ and the domain B_ϵ with $D(\mathfrak{x}_\epsilon) = d_\epsilon$. Lemma (b) of §4 excludes degeneration of B_ϵ because of the assumption 2) with fixed positive α. Hence using also Lemma 6 of §2 we can choose at least a subsequence ϵ for which B_ϵ tends to a limiting circular domain B. Equicontinuity of the $\mathfrak{x}_\epsilon$ on the boundary circles follows exactly as before in No. 1 and therefore we have at least a subsequence of vectors $\mathfrak{x}_\epsilon$ uniformly

converging to a potential vector $\mathfrak{x}$ for which $d \leqq \underline{d}$ and which satisfies the conditions of the problem I for the boundary Γ and for the domain B. Our theorem is thus proved.

At the same time we recognize that $\mathfrak{x}$ defines a minimal surface and that we have proved the following

APPROXIMATION THEOREM. *Under the assumption $d < d'_\epsilon + d''_\epsilon - \alpha$ a subsequence at least of the minimal surfaces S_ϵ for the contours $\Gamma^{(\epsilon)}$ converges to a minimal surface solving Plateau's problem for the contour* Γ, *provided that the areas of the surfaces S_ϵ remain bounded.*[26]

Exactly as in §3, No. 4 for $k = 1$ our statement concerning semicontinuity of $\underline{d}$ can be replaced by a *Continuity Theorem* the formulation and the proof of which are identical to that in §3.

4. **Conformal mapping for doubly connected domains.** As an application of our theory for the special case of $m = 2$ dimensions of the $\mathfrak{x}$-space we obtain the result:

Every twofold connected domain G in an x, y-plane bounded by two Jordan curves Γ_1, Γ_2 can be mapped conformally on an annular ring B. The conformal mapping of the open domains G and B establishes a continuous one-to-one correspondence between the boundaries.

The mapping theorem and the continuity of the mapping of C on Γ follows from our general theory for $m = 2$ if we can verify a priori the inequality $d < d_1 + d_2$ where d_1 and d_2 are the minima for the corresponding minimum problems for the single contours Γ_1, Γ_2 resp. Since for single contours the problem was solved and thus the possibility of the mapping theorem was proved we know that d_1 and d_2 are the areas included in Γ_1 and Γ_2 respectively. If we knew that d also represents the area of G we would have immediately $d = d_1 - d_2$ if, as we shall assume, Γ_2 is included in Γ_1; and all the more $d < d_1 + d_2$. But the relation $d = d_1 - d_2$ follows only after we have solved problem I for $k = 2$ and thus have proved the possibility of the conformal mapping of B on G. Therefore it is now our task to verify the inequality $d < d_1 + d_2$ directly.

To this end we use a typical argument of continuity based on the theorems of No. 3. Suppose problem I *is* solved for a domain G, then the equality $d = d_1 - d_2$ and hence the inequality $d < d_1 + d_2$ holds. Now let G' be a neighbouring domain bounded by Γ'_1 and Γ'_2 and originating from G by a small deformation of the x, y-plane as described in the continuity theorem of No 3. Then by this theorem the inequality $d' < d'_1 + d'_2$ will be obtained also for the domain G'; this remark enables us, if we start with a circular domain G, for which the solvability of problem I and the conformal mapping is trivial, to pass with a finite number of steps to an arbitrary polygonal domain G. Precisely, we suppose that for all domains G in consideration the area d_1 included in the

[26] For $k = 2$ Douglas has shown that this theorem can be extended to cases of unbounded areas. Cf. Radò's report.

outer contour Γ_1 is equally bounded by $d_1 < M$, while all the inner contours Γ_2 shall include an area $d_2 > p$ where p is a fixed positive number. If for G problem I is solved we have certainly $d = d_1 - d_2$.

Now we consider a transformation of the x, y-plane into the x', y' plane: $x' = x + \xi(x, y)$, $y' = y + \eta(x, y)$, where the functions ξ, η are continuous in $G + \Gamma$ and the first derivatives of ξ, η are allowed to be discontinous only along straight lines in G. G', Γ', d', d_1', d_2' refer to the transformed domain. With a small positive ϵ and $\delta = (1 + 2\epsilon)^2 - 1$ we require

$$|\xi_x| < \epsilon, \quad |\xi_y| < \epsilon, \quad |\eta_x| < \epsilon, \quad |\eta_y| < \epsilon, \quad |\xi| < \epsilon, \quad |\eta| < \epsilon$$

and we choose ϵ so small that $d_1' < M$, $d_2' > p$, $\delta M < \frac{1}{2}p$ and $|d_1 - d_1'| < \frac{1}{2}p$. To the straight lines of discontinuity in G correspond analytic lines in B by the supposed conformal mapping and therefore our continuity theorem can be applied. It states for the lower limits in problem I for G'

$$d' < d(1 + 2\epsilon)^2 = d + d\delta < d + M\delta < d + \tfrac{1}{2}p,$$

hence, since according to our assumption, $d = d_1 - d_2 < d_1 - p$, we have $d' < d_1 - p + \frac{1}{2}p < d_1' < d_1' + d_2'$. Therefore for G' the inequality $d' < d_1' + d_2'$ is proved; consequently problem I can be solved and d' is the area of G'. Thus the inequality $d' < d_1' + d_2'$ implies again the much stronger relations

$$d' = d_1' - d_2' < d_1' + d_2' - 2p.$$

We apply this result by joining an arbitrary polygonal domain $G = G_N$ with a circular domain G_0 by a sequence of domains G_1, G_2, $\cdots$ with the following properties: a) $G_1, \cdots G_N$ are bounded by polygons, b) for all these domains we have the fixed bounds M, p described above, c) G_r is transformed into G_{r+1} by a transformation of the type described above.[26a]

Since for G_0 problem I is solved by a domain B congruent to G_0, the solvability follows successively for $G_1, G_2, \cdots, G_N = G$. By a passage to a limit with polygonal domains G we finally can solve the problem, on the basis of the approximation theorem of No. 3, for any domain G bounded by arbitrary Jordan curves. Therefore the possibility of the conformal mapping of the interior of B on that of G is established and at the same time the continuous mapping of C on Γ. That also the mapping of Γ on C is a continuous correspondence is seen exactly as in the case of one contour in §3.

Incidentally, we now can show for arbitrary m in exactly the same manner as in the case $k = 1$ that: *Our solution of Plateau's problem establishes a one-to-one correspondence between the boundaries C and Γ.*

[26a] The possibility of constructing such a chain is obvious on the basis of elementary topology. E.g. we may assume that two consecutive polygonal domains differ only in one corner. The transformations of one polygon into the next one can be effected by piecewise linear transformations.

6. Plateau's Problem for Arbitrary k

In the case $k > 2$ we may ease our task by assuming that Plateau's Problem is already solved for smaller values of k and further that the lower semicontinuity of the lower limit $\underline{d}$ in its dependence on the boundary Γ is already established for smaller values of k.

Also the possibility of conformal mapping of any domain B of the u, v-plane on a circular domain can be assumed provided that B is bounded by less than k Jordan curves. This mapping implies a one-to-one correspondence also on the boundaries. Therefore it follows from the initial remarks in §2 that in variational problems I for less than k contours it does not matter whether we choose B as a circular domain or as any domain with a corresponding number of Jordan contours.

In the Problem I we may take C_1 as the unit circle, C_2 as a concentric circle inside of C_1 and the other circles C_ν in the annular ring between C_1 and C_2. Again let B_n, $\mathfrak{x}_n$ be a minimizing sequence of such circular domains and corresponding potential vectors. Again we make the assumption

$$d < d' + d'' \tag{21}$$

for every partition of C in $C' + C''$ and correspondingly of Γ in $\Gamma' + \Gamma''$ (cf. notations of §4).

Under this condition we shall show: *The problem* I *has a solution which represents a minimal surface bounded by* Γ *and conformally equivalent to a circular domain* B. Further: *The theorems of* §5 *concerning lower semicontinuity and continuity of the lower bound* d *with respect to the boundary, the approximation theorem, the theorem on the possibility of conformal mapping of any plane domain bounded by* k *Jordan curves on a circular domain, and the one-to-one correspondence of the boundaries, subsist also for* $k > 2$.

1. Solution of the variational problem.

We state:

a) The sequence of domains B_n—or a suitable subsequence—converges to a circular domain B.

b) On each boundary circle of B the $\mathfrak{x}_n$ are equicontinuous functions of the arc length or the angle.

c) The potential vectors $\mathfrak{x}_n$—or a subsequence—converge uniformly in B to a potential vector $\mathfrak{x}$ which solves Problem I.

To prove a) we have to exclude the following possibilities:

First: B tends to separation in the sense explained in §4, in which case Lemma b) of §4 would yield $d = \lim d(B_n) \geqq d' + d''$ which is incompatible with (21).

Second: Two boundary circles of B_n, none of them shrinking to a point, become arbitrarily near each other for certain increasing indices n. This contradicts Lemma c) of §4.

Third: Certain boundary circles of B_n shrink to a point on one boundary

circle e.g. C_1 whose radius remains above a positive bound and which therefore—if we choose a suitable subsequence—converges to a limiting circle.

To exclude this last type of irregularity is a somewhat more delicate task. We illustrate our reasoning by the figure 5 for the typical case where the circle C_3 converges to a point R on the unit circle C_1. Around R we again introduce polar coordinates r, ϑ. On account of Lemma 5 of §2 we see: For an arbitrarily small fixed δ—and sufficiently large n—there exists a circular arc c of radius r with $\delta \leqq r \leqq \sqrt{\delta}$ around R joining two points P, Q on C_1 and including C_3, on which the oscillation of the vector $\mathfrak{x} = \mathfrak{x}_n$ is less than $[2\pi M/|\log \delta|]^{\frac{1}{2}} = \epsilon(\delta)$, M being a bound of the values $D(\mathfrak{x}_n)$. Without loss of generality we may assume that on the arc c we have $|\mathfrak{x}_n| < \epsilon(\delta)$ for all n.

The arc c together with the arc b': PRQ of C_1 forms a closed curve $c + b'$ and together with the complementary arc b'' of C_1 forms the closed curve $c + b''$. If R is a point of equicontinuity on C_1 then the oscillation of our vector $\mathfrak{x}$ on the curve $T' = c + b'$ remains less than a quantity which tends to zero with δ. If R is a point of nonequicontinuity the same is true for the curve $T'' = c + b''$

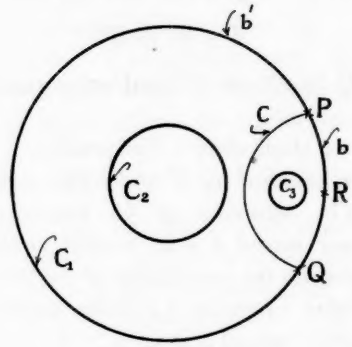


Fig. 5

on account of Lemma 6 §2 at least after we choose a proper subsequence $\mathfrak{x}_n$. The image of the arc b' by the representation $\mathfrak{x}_n(u, v)$ is called β', that of b'' is called β'' so that we have $\beta' + \beta'' = \Gamma_1$. One of these arcs β' or β'' has a diameter tending to zero with δ, which also is true for the diameter of the image γ of c.

c divides $B = B_n$ into the domains B' and B'', with $B' + B'' = B$ and with connectivities $k' < k$ and $k'' < k$ so that our theory is established for them according to our assumption. The boundaries of B' are mapped by $\mathfrak{x}_n$ on a system of continuous curves Γ'^{δ} and the boundaries of B'' on another system of such curves Γ''^{δ}. If $\delta = \delta_n$ tends to zero, Γ'^{δ} and Γ''^{δ} tend to a partition of Γ into Γ' and Γ'' plus an isolated point $\mathfrak{x} = 0$ which however according to Lemma 2 of §2 will not affect the lower limit of the Dirichlet Integral.

Now in the minimum problem defining $d(B) = d(B_n)$ we ease the conditions for admission by allowing $\mathfrak{x}$ to be discontinuous on c but only in such a way that the boundaries of B' and B'' remain mapped on Γ'^{δ} and Γ''^{δ}. Then the lower limit certainly will become smaller or at most equal to the value $d(B)$. The

new minimum problem refers obviously to two independent domains B' and B'' and the corresponding lower limit is equal to the sum of the two lower limits for B' and B'' under the condition that their boundaries are mapped on Γ'^{δ} and Γ''^{δ}. Thus in an obvious denotation we have

$$d(B) \geqq d(B', \Gamma'^{\delta}) + d(B'', \Gamma''^{\delta}),$$

hence all the more

$$d(B) \geqq d(\Gamma'^{\delta}) + d(\Gamma''^{\delta})$$

if on the right side we introduce the absolute lower bounds corresponding to the boundaries Γ'^{δ} and Γ''^{δ}. Now we let n tend to infinity and at the same time δ to zero; then the left side of our inequality tends to $\underline{d}$ while on account of the semicontinuity of $\underline{d}$ for $k' < k$ and $k'' < k$ the lower limit of the right side can not be less than $d' + d''$. Therefore we obtain immediately $d \geqq d' + d''$ which contradicts our condition $d < d' + d''$. Hence the third possibility of degeneration of our minimizing sequence B_n is excluded.

Therefore the sequence of domains B_n, or at least a subsequence, converges to a circular domain B bounded by k circles. The equicontinuity of the boundary values of the vectors $\mathfrak{x}_n$ now follows exactly as in §6 and also exactly in the same way we obtain an admissible limiting potential vector $\mathfrak{x}$ in B for which $D(\mathfrak{x}) = d$. This potential vector therefore solves Problem I defining a surface S in the $\mathfrak{x}$-space.

2. The solution S is a minimal surface.

To identify the potential surface S as a minimal surface requires somewhat deeper reasoning for $k > 2$. We first establish the natural boundary conditions on C_ν originating from the degree of freedom in the boundary representation for the curves Γ_ν. For that we consider one of the boundary circles e.g. C_1 (C_1 may include the others). Let r, ϑ again be concentric polar coordinates and let $\lambda(r, \vartheta)$ be a function with continuous derivatives in $B + C$, vanishing identically except in a neighboring ring adjacent to C_1. Then the natural boundary condition for C_1 is exactly as before in §§4, 5 uniformly in λ:

$$\lim_{r\to1} \int_0^{2\pi} \lambda(r, \vartheta) r\mathfrak{x}_r\mathfrak{x}_\vartheta \, d\vartheta = 0.$$

Or if we put as before

$$r\mathfrak{x}_r\mathfrak{x}_\vartheta = \Im(w^2\varphi(w)) = p(u, v) = p(r, \vartheta),$$

$p(u, v)$ being a potential function,

$$\lim_{r\to1} \int_0^{2\pi} \lambda(r, \vartheta) p(r, \vartheta) \, d\vartheta = 0. \tag{22}$$

If Q is, as in §3, a point in B we may again choose λ in the strip adjacent to C_1 as the Poisson kernel belonging to the interior of the circle with the radius r and the point Q. Then

$$\int_0^{2\pi} \lambda(\rho, \vartheta)p(\rho, \vartheta)\, d\vartheta = \Pi_\rho(Q) = \Pi_\rho(r, \vartheta)$$

is a potential function of u, v or r, ϑ regular for $r < \rho$ and equal to the fixed potential function p on $r = \rho$. Now formula (22) states that $\Pi_\rho(Q)$ tends to zero uniformly for every closed subdomain as ρ tends to 1. From this there follows by an elementary reasoning:

The potential function p has the boundary values zero on C_1.

PROOF. The difference $p(Q) - \Pi_\rho(Q) = \omega_\rho(Q) = \omega_\rho(r, \vartheta)$ is a potential function which has the boundary values zero for $r = \rho$ and exists in a ring $l \leqq r \leqq \rho$ with fixed l which lies entirely in B. On $r = l$ the function ω_ρ is bounded because p is regular in B, and for $\rho \to 1$ we have $\Pi_\rho(Q) \to 0$. Now the potential function $\omega_\rho(Q)$ can be analytically extended by reflection beyond $r = \rho$ into the ring $l \leqq r = \rho^2/l$. Therefore in every interior concentric ring R the derivatives of ω_ρ are equally bounded with the same bound N. Now if Q in B has a distance less than h to C_1 and all the more to C_ρ then we certainly have for sufficiently small h, $|\omega_\rho(Q)| < Nh$. Therefore for $\rho \to 1$ because of $\omega_\rho \to p$ we obtain, for Q fixed, $|p(Q)| \leqq Nh$, which proves that $p(Q)$ indeed approaches zero uniformly when the distance h to C_1 tends to zero.

Therefore p is *regular also on* C_1 and assumes there the values zero.

In the same way we show: If α_ν is the center of the circle C_ν of radius ρ_ν then the imaginary part of the function $(w - \alpha_\nu)^2 \varphi(w)$ vanishes on C_ν.

$$\Im[(w - \alpha_\nu)^2\varphi(w)] = 0 \text{ on } C_\nu. \tag{23}$$

These boundary conditions which for $k = 2$ and $k = 1$, $\alpha_\nu = 0$ amount to the fact that $\Im(w^2\varphi(w)) = 0$ throughout B (not only at the boundaries) have here a more complex appearance.

Our task is now to obtain the desired relation $\varphi(w) = 0$ in B from these conditions (23) by combining them with the further conditions which express the fact that variations of the size and the position of the boundary circles C_ν cannot improve the lower limit $\underline{d}$.

The variation of the radii of the circles C_ν yields in exactly the same way as in §6 the conditions

$$\Re \int_{C'_\nu} (w - \alpha_\nu)^2 \varphi(w)\, d\vartheta = 0 \tag{24}$$

where this integral is extended over a circle C'_ν concentric to C_ν and sufficiently near to it.

We shall show by a similar reasoning that the variation of the center α_ν of C_ν yields for C_ν the two additional conditions

$$\Re \int_{C'_\nu} (w - \alpha_\nu)\varphi(w)\, d\vartheta = 0 \tag{25}$$

$$\Im \int_{C'_\nu} (w - \alpha_\nu)\varphi(w)\, d\vartheta = 0, \tag{26}$$

where again the integrals are extended over concentric circles C'_ν sufficiently near to C_ν.

To establish these variational conditions we perform as in §6 a simple simultaneous variation of the minimal vector $\mathfrak{x}$ and the minimal domain B. First we replace B by another domain B' which is bounded by circles C'_ν concentric and sufficiently close to C_ν. Let R_ν denote the ring between C and C'. The values of the minimizing potential vector $\mathfrak{x}$ on C'_ν are analytic functions $j_\nu(\vartheta)$ of the polar angle corresponding to this circle.

We now replace B' by another circular domain B^* depending on a parameter ϵ with boundaries C^*_ν which originate from the boundaries C'_ν by parallel transformations. We further transfer the boundary values $f_\nu(\vartheta)$ from C'_ν to C^*_ν by the corresponding parallel transformation. With these fixed boundary values we solve the boundary value problem for B^* and thus define in B^* a potential vector

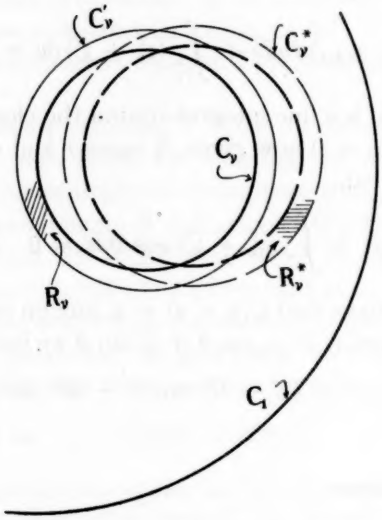


FIG. 6

$\mathfrak{z}(u, v, \epsilon)$ whereby $\mathfrak{z}(u, v, o) = \mathfrak{x}(u, v)$ is the original minimizing vector. Affixing the ring R_ν to C^*_ν by a parallel transformation we obtain a domain $B_\epsilon = B^* + \sum_\nu R^*_\nu$, R^*_ν denoting the ring R_ν after the parallel transformation. In B_ϵ we define a vector $\mathfrak{z}$ which is equal to $\mathfrak{z}(u, v, \epsilon)$ in B^* and which in the ring R^*_ν has the same values as $\mathfrak{x}$ in the corresponding points of the ring R_ν. Obviously the vector $\mathfrak{z}$ and the domain B_ϵ are admissible in our Problem I and hence because of the minimum property of B and $\mathfrak{x}$ we have

$$\begin{aligned} D_{B_\epsilon}(\mathfrak{z}) &= D_{B^*}(\mathfrak{z}) + \sum_\nu D_{R^*_\nu}(\mathfrak{z}) \\ &= D_{B^*}(\mathfrak{z}) + \sum_\nu D_{R_\nu}(\mathfrak{x}) \\ &\geqq D_B(\mathfrak{x}) = D_{B'}(\mathfrak{x}) + \sum_\nu D_{R_\nu}(\mathfrak{x}). \end{aligned}$$

Therefore we obtain $D_{B^*}(\mathfrak{z}) \geqq D_{B'}(\mathfrak{x})$. In other words: The minimum property of $\mathfrak{x}$ with respect to the position of the centers subsists also for subdomains bounded by circles concentric to the boundary circles C_ν.

After this preparation we again can take advantage of the analytic character of the boundary values of $\mathfrak{x}$ on C'_ν. We suppose now that B^* originates from B' by shifting only one circle C'_ν parallel to the u-axis by the distance ϵ, the other circles remaining fixed. Then $\mathfrak{z}\,(u, v, \epsilon)$ has continuous derivatives of every order with respect to u, v, ϵ[27] in B^* and on its boundary and from our result it follows that

$$\frac{\partial}{\partial \epsilon} D_{B^*}(\mathfrak{z}) = 0 \quad \text{for } \epsilon = 0.$$

Again applying elementary rules for the differentiation of an integral with respect to a parameter we get

$$2 \iint_{B'} (\mathfrak{z}_u \mathfrak{z}_{u\epsilon} + \mathfrak{z}_v \mathfrak{z}_{v\epsilon})\, du\, dv + \int_{C^*_\nu} (\mathfrak{z}_u^2 + \mathfrak{z}_v^2)\, dv = 0 \quad \text{for } \epsilon = 0$$

where the second integral is a line integral around the closed curve C^*_ν. Green's theorem on account of $\Delta \mathfrak{z} = 0$ now gives, if again r and ϑ are polar coordinates concentric to C_ν and $s = r\vartheta$,

$$-2 \int_{C^*_\nu} \mathfrak{z}_\epsilon \mathfrak{z}_r\, ds + \int_{C^*_\nu} (\mathfrak{z}_u^2 + \mathfrak{z}_v^2) \cos \vartheta\, ds = 0 \quad \text{for } \epsilon = 0.$$

Now on C^*_ν for $\epsilon = 0$ we have that $\mathfrak{z}_r(u, v, o) = \mathfrak{x}_r$ and on account of the definition $\mathfrak{z}_\epsilon(u, v, o) = \mathfrak{x}_u$. But from $\mathfrak{x}_r = \mathfrak{x}_u \cos \vartheta + \mathfrak{x}_v \sin \vartheta$ we infer

$$-(\mathfrak{x}_u^2 + \mathfrak{x}_v^2) \cos \vartheta + 2\mathfrak{x}_u \mathfrak{x}_r = (E - G) \cos \vartheta + 2F \sin \vartheta$$
$$= (1/r)\, \Re[(w - \alpha_\nu)\varphi(w)].$$

Hence our equation becomes

$$\Re \int_{C'_\nu} (w - \alpha_\nu)\, \varphi(w)\, d\vartheta = 0. \tag{25}$$

In the same way we find by performing a variation (parallel transformation) of C' parallel to the v-axis that

$$\Im \int_{C'_\nu} (w - \alpha_\nu)\varphi(w)\, d\vartheta = 0. \tag{26}$$

The totality of our variational conditions can be written in a concise form. First (25) and (26) can be combined in the form

$$\int_{C'_\nu} (w - \alpha_\nu)\varphi(w)\, d\vartheta = 0. \tag{27}$$

[27] To verify this we only need the Poisson kernel for a single circle C_ν.

Further, on C'_ν we have $w - \alpha_\nu = re^{i\vartheta}$ and $i(w - \alpha_\nu)\, d\vartheta = dw$. This last condition therefore can be written in the form

$$\int_{C'_\nu} \varphi(w)\, dw = 0. \tag{28}$$

Similarly we integrate the condition (23) for C_ν with respect to ϑ and combine this equation with (24). We thus obtain

$$\int_{C'_\nu} (w - \alpha_\nu)\varphi(w)\, dw = 0$$

and on account of (28)[28]

$$\int_{C'_\nu} w\varphi(w)\, dw = 0. \tag{30}$$

It remains to prove the THEOREM. *If in a circular domain B the conditions (23), (28), (30) hold, then φ vanishes identically in B.*[29]

With the proof of this theorem Plateau's Problem for k contours is solved.

3. **Proof of the preceding theorem.**[30] We shall show that the assumption that $\varphi(w)$ is not identically zero leads, by counting the zeros of $\varphi(w)$ in B, to an absurd consequence. Accordingly let us suppose that φ does not vanish identically. Then in the closed domain $B + C$ φ has only a finite number of zeros, and consequently the function $(w - \alpha_\nu)^2 \varphi(w)$ which is real on C_ν also has only a finite number of zeros on C_ν. We first prove that at C_ν this function has at least four zeros which separate arcs where the function is positive from those where it is negative. Without restricting the generality we may assume that the circle under consideration is the unit circle, so that for $w = c^{i\vartheta}$ and $dw = iwd\vartheta$,

$$w^2\varphi(w) = f(\vartheta)$$

[28] Cauchy's integral theorem now shows that the integrals of the functions $\varphi(w)$ and $w\varphi(w)$ not only vanish for these closed curves C'_ν, but for every closed curve lying in the domain $B + C$. That means first that the function φ is the derivative χ' of a single valued analytic function $\chi(w)$ in $B + C$ and that again χ is the derivative of another single valued regular analytic function $\psi(w)$ in $B + C$. In other words, we have

$$\varphi(w) = \psi''(w)$$

where $\psi(w)$ is a single valued regular analytic function in $B + C$.

This statement concentrates the essence of our variational conditions with respect to the domain B. Thus our variational conditions present themselves in the following form

$$\mathfrak{J}[(w - \alpha_\nu)^2\psi''(w)] = 0 \text{ on } C_\nu \qquad \varphi = \psi''(w) \tag{29}$$

with ψ single valued and regular in $B + C$.

[29] We can express our condition in the following form: Let $\partial/\partial r$ denote the normal derivative on the circle C: then $(w - \alpha_\nu)^2\psi''(w) = \partial^2\psi/\partial r^2$ is the second normal derivative of ψ on C_ν and (29) expresses that it has a vanishing imaginary part on C_ν. From this we want to infer: ψ is a linear function of the form $\psi = a + bw$.

[30] The method of the following proof was kindly suggested to me by Hans Lewy in Berkeley.

is a real periodic function of the angle ϑ for which, by (30) and (28),

$$\int_0^{2\pi} f(\vartheta)\, d\vartheta = 0 \tag{31}$$

$$\int_0^{2\pi} f(\vartheta) \cos(\vartheta - \vartheta_0)\, d\vartheta = 0 \tag{32}$$

holds for every ϑ_0. But such a periodic function $f(\vartheta)$ has indeed at least 4 zeros as described above.[31]

Now let $N \geqq 0$ be the number of zeros of $\varphi(w)$ in the interior of B. This number is expressed, if the circle C_1 contains the other circles C_ν, by the formula

$$2\pi i N = \int_{C_1'} d \log \varphi + \sum_{\nu=2}^{k} \int_{C_\nu'} d \log \varphi(w)$$

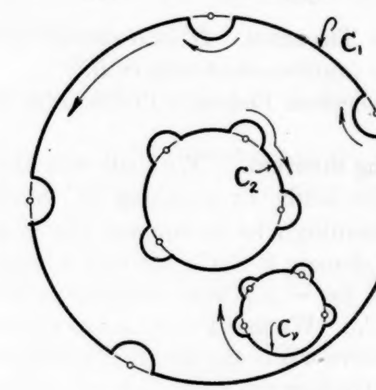


FIG. 7

where now, C_ν' denotes a path of integration coincident with C_ν except for little halfcircles circumventing the zeros of φ on C_ν as in figure 7. But

$$2\pi i N = \int_{C_1'} d \log [(w - \alpha_1)^2 \varphi(w)] - \int_{C_1'} d \log (w - \alpha_1)^2$$
$$+ \sum_{\nu=2}^{k} \left\{ \int_{C_\nu'} d \log [(w - \alpha_\nu)^2 \varphi(w)] - \int_{C_\nu'} d \log (w - \alpha_\nu)^2 \right\}$$

or

$$2\pi i N = \int_{C_1'} d \log [(w - \alpha_1)^2 \varphi(w)] - 4\pi i + \sum_{\nu=2}^{k} \left\{ \int_{C_\nu'} d \log [(w - \alpha_\nu)^2 \varphi(w)] + 4\pi i \right\}. \tag{33}$$

[31] The equation (31) shows that there are at least two such intervals $f(\vartheta) > 0$ and $f(\vartheta) < 0$ separated by 2 zeros. If there were not at least four such intervals with alternatingly different signs (their number must be even) with at least 4 zeros as endpoints, that is, if we only had two zeros we could by changing the origin $\vartheta = 0$ suppose these zeros in the form β and $-\beta$. Then the function $(\cos \vartheta - \cos \beta) f(\vartheta)$ would have the same sign for all values ϑ and the integral $\int_0^{2\pi} (\cos \vartheta - \cos \beta)\, f(\vartheta) d\vartheta$ could not vanish, in contradiction to (32).

But each of the small halfcircles circumventing a zero winds around this zero in the negative sense as indicated in figure 7 and leads from real values of the integrand $(w - \alpha_\nu)^2\varphi(w)$ to real values and therefore contributes to the integral a multiple of $-\pi i$ while the other part of C'_ν on which $(w - \alpha_\nu)^2\varphi(w)$ is real does not contribute anything to the imaginary part. Each of the integrals therefore contributes at least $-4\pi i$ to the imaginary part; hence the right side of (33) is, except for the factor i, less than or equal to -8π and it would follow that

$$0 \leqq 2\pi N \leqq -8\pi,$$

which is absurd. Consequently we have $\varphi(w) = 0$ identically and our theorem is proved.

4. **Additional remarks.** a) Exactly in the same way as in the previous paragraphs by applying our result to the case $m = 2$ we obtain the following *theorem on conformal mapping*:

Every domain of the x, y-plane bounded by k Jordan curves can be mapped conformally on a circular domain of the u, v-plane.[32] *The mapping implies a one-to-one correspondence of the boundaries.*

b) Our solution of Plateau's Problem furnishes also for $m > 2$ a *one-to-one correspondence of the boundaries C and Γ.* The proof proceeds exactly on the same lines as in §3.

c) The lower limit depends on the boundary Γ in a *lower semicontinuous way.* Also this fact follows exactly along the same lines as in the case $k = 2$.

d) The approximation theorem and the continuity theorem of §§3, 5 remain valid for $k > 2$.

Thus the statements in the introduction of §6, assumed as true for $k - 1$ contours, are established for k contours and hence the proof is completed.

Finally we notice: *the lower limit d is equal to the area of the minimal surface S.*

7. Solution of Plateau's Problem and the Problem of Least Area Based on Conformal Mapping

1. **Preliminary remarks.** If instead of basing the proof of Riemann's mapping theorem and its generalizations on the general theory of Plateau's Problem we do not decline taking advantage of fundamental facts concerning conformal mapping of domains in a plane we can considerably simplify the solution of Plateau's Problem.[33]

Thus we also can easily establish the fact: *Our solution of Plateau's Problem furnishes at the same time the surface of least area bounded by Γ.*

[32] This theorem for arbitrary k has previously been proved in different ways by Koebe and by the author (cf. Hurwitz-Courant, *Funktionentheorie*, Berlin 1931). The proof contained in the present paper seems to be essentially more elementary.

[33] For these facts cf. Hurwitz-Courant, *Funktionentheorie*, Part III, Berlin, 1931, Courant, Crelle's Journ. Vol. 165 (1931) p. 247 ff. and a paper to be published.

We again start by solving Problem I as before. Then the use of conformal mapping serves to identify the solution as a minimal surface.[34]

The domain B can either be chosen as a circular domain as before or—in a way more elementary from the point of view of conformal mapping—as a "Schlitz" domain. A *"Schlitz" domain* is the whole u, v-plane including the point at infinity, but cut along k straight segments, all parallel to the u-axis. The two edges of each of the cuts form a component C_ν of the boundary C.

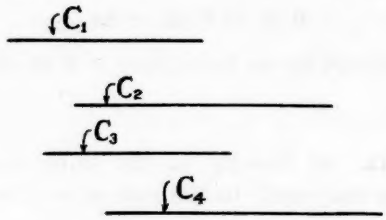


FIG. 8

The new method for the solution of Plateau's Problem is based upon the two following theorems:

THEOREM (α). *Every k-fold connected domain of the plane with k simple Jordan contours can be mapped conformally on a domain B e.g. of the "Schlitz"-type B whereby the boundaries are mapped in a one-to-one continuous way.*[35]

THEOREM (β). *Let a k-fold connected domain G in the u, v-plane consist of two or more parts G_1, G_2, $\cdots$ which may be separated or adjacent and which are connected in the following way: Some analytic arc of the boundary of a part G_ν shall correspond by an analytic transformation of $w = u + iv$ to an analytic arc of the boundary of (another or the same) G_ν and corresponding points may be considered as identical. (Thus interior points of these arcs become interior points of G.) Then the domain G can be mapped conformally on a simple domain B of the above type so that corresponding points on the corresponding parts of the boundaries of the G_ν become identical and that the boundary of G corresponds in a one-to-one way to that of B. In this it is assumed that such a mapping is topologically possible.*[36]

In other words: Theorem β says that a domain G connected *in abstracto* can be realized *in concreto* geometrically by conformal mapping of the composing domains.

An example for a domain G as considered in our theorem is the unit circle in the u, v-plane, cut along the v-axis between $v = 0$ and $v = \frac{1}{2}$, the two edges

[34] Another method using conformal mapping was as mentioned before, pointed out by Douglas and Radò and in a somewhat different way by McShane.

[35] The one-to-one correspondence of the boundaries has to be interpreted in such a way that the opposite points at a cut have to be counted as different.

[36] This theorem is also of a quite elementary character although apparently never stated explicitly. I shall give a simple proof for (α) and (β) and for some generalizations on piecewise Riemannian manifolds in a separate paper.

being coordinated by the transformation $v' = 2v^2$. Theorem β states: The unit circle with this cut can be mapped on the unit circle G' with another cut so that after the mapping corresponding points become coincident.

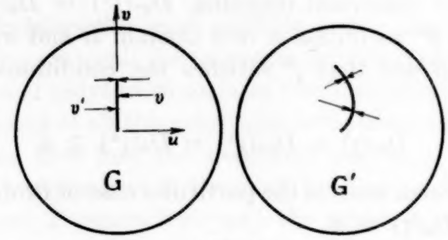


FIG. 9

From the point of view of method it may be worthwhile mentioning that (α) and also (β) can be obtained directly by nearly the same reasoning as in §§2–6 only on the basis of the boundary value problem of potential theory and the corresponding minimum property of the Dirichlet Integral. By this remark the subsequent proof of the least area property could be made independent of preliminary knowledge in conformal mapping.

2. **Solution of Plateau's problem.** Suppose Problem I is solved by the vector $\mathfrak{x}$ in the domain B. Our task is to prove $\varphi(w) = 0$. We cut B by an arbitrary straight segment L through a point P; e.g. a segment $v = \text{const}$.

Thus we obtain a new domain B' with the boundary C' consisting of C and the two edges of the cut L. Between the two edges of L we establish an analytic one-to-one correspondence leaving the endpoints of L intact. (In other words we define a domain G where corresponding points on the two edges of L are identified.)

Now we replace Problem I by the following:

PROBLEM II. *Let $\mathfrak{z}$ be a vector in B' continuous in B' and its boundary C', mapping C on Γ, having continuous boundary values identical at corresponding points of the two edges of the cut L and having piecewise continuous first derivatives in B'. Then we ask for the minimum value d_0 of $D_{B'}(\mathfrak{z})$ obtained by admitting all possible correspondences between the edges of L.*

In other words: we renounce in Problem I continuity of $\mathfrak{x}$ along L, but we still require that $\mathfrak{x}$ map the two edges of L on the same curve in the $\mathfrak{x}$-space, the mappings of the two edges being equivalent by an analytic transformation. Now we state the following

THEOREM: *The minimum d_0 in Problem* II *is the same as in Problem* I *and is attained by the original solution $\mathfrak{x}$ of Problem* I. This follows easily by means of our theorems (α) and (β).

We consider any vector $\mathfrak{z}$ corresponding to a domain B' complying with the conditions of Problem II for our domain B', and defining along the cut L a one-to-one analytic correspondence. On account of the theorems (α) and (β)

we can map B' conformally on a new domain B^* so that corresponding points of the edges of L become identical. Let the vector $\mathfrak{z}$ pass by this conformal mapping into the transformed vector $\mathfrak{z}^*$. We certainly have, by the invariance of the integral under conformal mapping, $D_{B^*}(\mathfrak{z}^*) = D_{B'}(\mathfrak{z})$. By adding the image points of L to B^* we obtain a new domain B and we are now sure that $\mathfrak{z}^*$ is continuous in B and that $\mathfrak{z}^*$ satisfies the conditions of Problem I for B. Hence

$$D_{B'}(\mathfrak{z}) = D_{B^*}(\mathfrak{z}^*) = D_B(\mathfrak{z}^*) \geqq d$$

which proves our theorem since in the particular case of Problem I for the original domain B we have $D_B(\mathfrak{x}) = d$.

Problem II gives us the advantage of more freedom for the variation than we had in Problem I. We exploit the minimum property of the solution $\mathfrak{x}$ of Problem I with respect to Problem II in the following way: Assuming our cut L through P as a segment $v =$ constant we replace $\mathfrak{x}(u, v)$ by $\mathfrak{z}(u, v) = \mathfrak{x}(u + \epsilon\lambda(u, v), v)$ where ϵ is a small parameter and $\lambda(u, v)$ vanishes identically everywhere except in a small rectangle R adjacent to one edge of L; λ is analytic, otherwise arbitrary in R, e.g. $\lambda = \lambda(u)\mu(v)$ where $\lambda(u) = 0$ at the endpoints of L and $\mu(v) = 0$ at the upper edge of R.

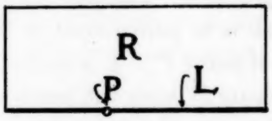


FIG. 10

Now we have $\mathfrak{z}_\epsilon(u, v, o) = \lambda\mathfrak{x}_u$ and the variational equation

$$\frac{\partial}{\partial\epsilon} D(\mathfrak{z}) = 0 \quad \text{for } \epsilon = 0$$

becomes, since $\mathfrak{z}$ and D depend analytically on ϵ, and differentiation may be performed under the integral sign

$$\int\int_R (\mathfrak{z}_u \mathfrak{z}_{u\epsilon} + \mathfrak{z}_v \mathfrak{z}_{v\epsilon})\, du dv = 0 \quad \text{for } \epsilon = 0,$$

or on account of Green's formula

$$\int_L \lambda \mathfrak{x}_u \mathfrak{x}_v \, du = 0.$$

Because of the arbitrariness of λ we obtain on L, hence in P, $\mathfrak{x}_u\mathfrak{x}_v = F = 0$.

If instead we take L as a segment $u - v =$ const. we get exactly in the same way

$$(\mathfrak{x}_u - \mathfrak{x}_v)(\mathfrak{x}_u + \mathfrak{x}_v) = E - G = 0$$

whereby our task of identifying our solution as a minimal surface is completed.

The preceding reasoning shows in general: *If the conditions for admission in*

Problem I *are lessened by admitting discontinuities of* $\mathfrak{x}$ *of the type described above, along a finite number of analytic arcs, the minimum in this new Problem* II *remains equal to d and is attained by the solution* $\mathfrak{x}$ *of Problem* I.

3. **Our solution furnishes the least possible area.** To prove the least area property of our solution we recall the following definition of the area of a surface S: If Π_n is a sequence of polyhedral surfaces which approach S, then the lowest possible limit of the areas of all the polyhedra is the area of S.

Now let us suppose first that the boundary curves Γ_ν are polygons. We consider an arbitrary polyhedron Π bounded by Γ with triangular sides S_μ—which is no restriction of generality—and the edges Λ_l. We subdivide the domain B into a set of triangles B_μ the edges L_l of which are straight lines or arcs of circles so that the edges L_l and the triangles B_μ correspond topologically to Λ_l and S_μ.

Now we pass from Problem I with the lower limit d to Problem II which differs from Problem I in so far as the admitted vectors $\mathfrak{z}$ are allowed to be discontinuous along the edge L_l mapping both sides of the edge L_l on the same analytic curve and establishing an analytic correspondence between the two edges. Then we know from No. 1 that the lower limit in Problem II again is d.

Finally we replace Problem II by Problem III imposing on the vectors the additional condition that they shall map L_l on Λ_l. For the lower limit d^* of Problem III we certainly have $d^* > d$ because the range of competition in Problem III is narrowed with respect to that in Problem II. But the lower limit in Problem III simply is equal to the sum of the corresponding lower limits referring to the mapping of the triangles B_μ on triangular surfaces bounded by the same lines Λ_l as the plane triangle S_μ and the minimum is the area of the triangles S_μ. (The boundary values in this mapping are analytic on account of the principle of reflection.) Therefore d^* is exactly the area of the polyhedron Π while d is the area of the solution of Problem I. This proves our statement immediately.

If Γ does not consist of polygons it can be approximated by polygons, and on account of the definition of the area S as the lower limit of the areas of approximating polyhedra we obtain the desired general theorem immediately from the preceding result concerning the special case of polygonal boundaries.

With this result we are able to interpret our previous inequalities $d < d' + d''$ etc. as inequalities referring to lower limits of areas, which intuitively is more satisfactory than the original definition.

8. Plateau's Problem for One-Sided Minimal Surfaces and for Minimal Surfaces of Higher Genus

It may happen—as was emphasized by Douglas—that a smaller lower bound $\underline{d}$ for the Dirichlet Integral is obtained if surfaces of a higher topological structure bounded by Γ are admitted in our variational problems, e.g. one-sided surfaces of the type of a Moebius strip or surfaces of higher genus topologically not

equivalent to a simple domain B of the u, v-plane. In this §8 we will show briefly that our method can dispose also of these cases in full generality.

To this end we have to represent surfaces S with the boundary $\Gamma = \Gamma_1 + \cdots + \Gamma_k$ by vectors $\mathfrak{x}(u, v)$ in a domain B of the u, v-plane with the prescribed topological structure. Such a domain B can be obtained from a simple domain B^* by establishing one-to-one analytic correspondences between certain parts of the boundaries of B^* and by considering corresponding points as identical, that is, attaching to them the same values of the vectors $\mathfrak{x}$. E.g. if S and hence

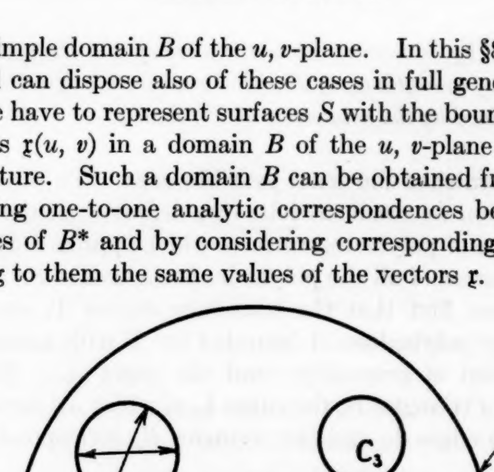


FIG. 11

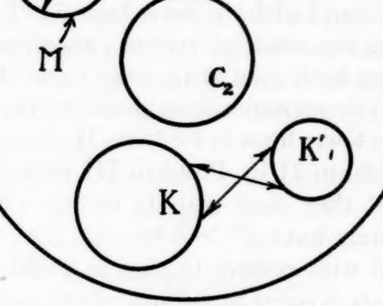


FIG. 12

B shall be an orientable (two sided) surface of genus p we may consider a circular domain B^* containing the point at infinity, bounded by k circles $C_1, \cdots C_k$—which correspond to Γ—and by p pairs K_ν, K'_ν of circles. K_ν shall correspond to K'_ν by a linear transformation of w which transforms B^* into another circular domain outside B^*. Corresponding points are identified and thus B^* becomes a domain of genus p.

Another suitable normal form of a plane domain of genus p with k boundary curves is the "*Schlitz domain.*" (Cf. §7). Such a Schlitz-domain consists of the whole u, v-plane cut by k straight segments $C_1, \cdots, C_k$ parallel to the

u-axis, further by p quadruples of such infinite cuts as in figure 12. Each quadruple of the infinite type of cuts consists of two pairs of cuts parallel to the u-axis extending to $u \to +\infty$ and lying vertically above the other, one pair separating the other pair. The edges of the cuts correspond to each other as indicated in figure 12 so that the corresponding points have the same coordinate u. Here also corresponding points are identified. Also limiting cases, when cuts of different pairs partly coincide must be considered.

Non-orientable (one-sided) surfaces can be topologically represented by plane domains in a similar way. E.g. we consider again a circular domain B^* with a boundary circle M which is transformed into itself by an analytical (linear) transformation T of the variable $\overline{w} = u - iv$ for which the iterated transformation $T\,T$ is the identity and which leaves no point on M invariant. Corresponding points on M shall again be identified; then B^* becomes one sided. A domain B bounded by k circles C_l, by p pairs K_ν, K'_ν of circles and in addition by q circles $M_1 \cdots M_q$ as described above is, for $q > 0$, a one sided surface of the *characteristic*[37] $N = 2p + q$. This characteristic number N and the character of orientability define the topological type of the surface.

In a similar way one can consider Schlitz domains B representing non-orientable surfaces of given characteristic.

A surface S in the $\mathfrak{x}$-space bounded by the fixed contours Γ can—as a limit of a sequence of such surfaces—degenerate into a surface of smaller characteristic or into two or more separate surfaces bounded by complementary sets Γ', Γ'', $\cdots$ of Γ respectively and having a sum of characteristics not exceeding N. If the prescribed topological structure is one sided, the degeneration may yield one or more one sided parts. Correspondingly we have to consider degeneration of the domain B in the u, v-plane. E.g. a sequence of circular domains B_n as above may tend to separation exactly as in §6, or one or more of the circles M_ν may shrink to points or the same may happen to a pair of circles K_ν, K'_ν., etc.

Similar degenerations may occur with Schlitz domains. We shall call such degenerate surfaces S or non-degenerate surfaces with lower characteristic with the total boundary Γ surfaces of a lower topological structure.

To investigate the possibility of a minimal surface with the boundary Γ and the prescribed topological structure we consider again the variational problem I for domains B of one of the types defined above, where the vectors admitted must have the same values in equivalent boundary points.

Exactly as in §4 we find the following result. *For a given boundary* Γ *we cannot improve the lower limit d in Problem* I *by passing to a lower topological type.*

Now the following general theorem holds:[38]

If to given Γ *and given topological structure the lower limit d in Problem* I*—or the corresponding area of a surface—is really smaller than for lower topological types,*

---

[37] The characteristic and the character of orientability are the only topological invariants of our surfaces except for the number k of contours. For one sided surfaces it is always possible to choose q either as 1 or as 2, or else to choose p as 0.

[38] See Douglas loc. cit. footnote 5.

then Problem I *has a solution* $\mathfrak{x}$ *which constitutes a minimal surface bounded by* Γ *having the prescribed topological structure and moreover giving the least possible area spanned by* Γ *in comparison to all surfaces of not higher topological type.*

To prove this theorem we first have to solve problem I along the lines of §6 excluding degeneration of the minimizing sequence B_n of domains and securing equicontinuity of the minimizing vectors by means of the supposed inequalities. Then we have to verify that the solution is a minimal surface, either by the method of §6 or by the method of §7, the latter of which requires a slight generalization of the mapping theorems of §7. The details, not essentially different from those in the preceding sections of this paper, will be given in another publication.

In the same way as for $N = 0$ the solution can be recognized as the solution of the problem of least area.

Concluding remarks. The methods developed in this paper permit applications to various other variational problems. At this place it may only be stated that a slight modification allows us also to solve Plateau's Problem if the boundaries of the minimal surface are not entirely fixed, but if parts of the boundary arcs are free to move on prescribed surfaces of less than m dimensions. In this case, under suitable assumptions, for a minimizing sequence of domains and vectors the existence of a limiting domain B again can be established. On the boundary C of the domain equicontinuity subsists on those parts which correspond to fixed parts of the boundary Γ, while for the other parts of the boundary C of B a certain uniformity in the approximation of the points $\mathfrak{x}(u, v)$ to points on the prescribed surfaces can be established whereafter the solution proceeds in the same way as in this paper. The method becomes particularly simple if the non fixed boundary surfaces are linear and therefore the principle of reflection can be applied to the construction of the solution.[39]

New York University.

[39] See a paper by E. Ritter soon to appear.

Annals of Mathematics
Vol. 38, No. 3, July, 1937

ARITHMETIC FUNCTIONS ON RINGS

By Morgan Ward

(Received September 1, 1936)

I. Introduction

1. The classic arithmetic properties of the Lucas[1] [1] function

$$u_n = \frac{\alpha^n - \beta^n}{\alpha - \beta}, \qquad \alpha + \beta \qquad \alpha\beta \quad \text{rational integers,}$$

can be shown to depend ultimately upon its periodicity to any integral modulus and its divisibility property: u_n divides u_m if n divides m. While the first property extends to any recurring series of integers[2], [3], [4] and the second may be similarly generalized,[5], [6], [7] the divisibility property is shared by numerous functions bearing no evident relation to recurring series, such as the totient function and its generalizations[8] and the polynomials γ_n of Halphen[9] associated with the rational multiplication of the Weierstrass $\wp$ and σ functions.

I show here that a much more extensive generalization is possible which not only reveals inner connections between the arithmetical properties of the Lucas function, but also appears to be of some independent interest.

The generalization consists in systematically developing the divisibility properties and modular periodicity of a function $y = \phi_x$ where x lies in an abstract commutative ring[2] while y lies in a structure[3] (usually another ring).

2. Let $\mathfrak{o}$ be a commutative ring, Σ a structure. We assume that to each element x of $\mathfrak{o}$ there corresponds a unique element

$$X = \phi(x) = \phi_x$$

of Σ. In the phraseology of general analysis, we call ϕ *a function on* $\mathfrak{o}$ *to* Σ. ϕ_a, ϕ_b are *values* of ϕ, a, b specific elements of $\mathfrak{o}$. We call ϕ *arithmetic* if ϕ_a divides ϕ_b in Σ whenever a divides b in $\mathfrak{o}$.

If $\mathfrak{o}$ and Σ are both the ring of rational integers, and $\phi_{-n} = -\phi_n$, an arithmetic function is equivalent to M. Hall's[6] divisibility sequence.[4] If $\mathfrak{o}$ itself is a structure, say a principal ideal ring, any homomorphism between $\mathfrak{o}$ and

---

[1] The [1] refers to the list of references concluding this paper.

[2] The Lucas function is brought within the scope of this generalization by letting $u_n = -u_n$.

[3] We use the term recently introduced by O. Ore[10] in these Annals. Equivalent terms are dual group, lattice.

[4] See the references in Hall[6] for earlier work on these sequences.

the values of ϕ in Σ with respect to cross-cut or union (O. Ore[10] pp. 416–419) defines an arithmetic function.[5]

3. We use the following notation: $\tilde{\mathfrak{o}}$ denotes the structure of all ideals of $\mathfrak{o}$, $a, \cdots, z$ elements of $\mathfrak{o}$, $\mathfrak{a}, \cdots, \mathfrak{z}$ ideals of $\mathfrak{o}$, $(a), \cdots, (z)$ principal ideals. We use $x \mid y$, $\mathfrak{x} \supseteq \mathfrak{y}$, $\mathfrak{y} \subseteq \mathfrak{x}$ for division in $\mathfrak{o}$ and $\tilde{\mathfrak{o}}$, xy, $\mathfrak{x}\mathfrak{y}$ for product, and $(\mathfrak{x}, \mathfrak{y})$, $[\mathfrak{x}, \mathfrak{y}]$ for union and cross-cut respectively.[6]

Corresponding entities in Σ are denoted by capital letters. We use $A, \cdots, Z$ for values of ϕ in Σ. These are on occasion imbedded in a ring $\mathfrak{O}$. Σ then denotes the structure of all ideals of $\mathfrak{O}$. We use $\mathfrak{A}, \cdots \mathfrak{Z}$ for elements of Σ and write $\mathfrak{X} \supseteq \mathfrak{Y}$, $\mathfrak{Y} \subseteq \mathfrak{X}$ for $\mathfrak{X}$ divides $\mathfrak{Y}$, $(\mathfrak{X}, \mathfrak{Y})$, $[\mathfrak{X}, \mathfrak{Y}]$ for union and cross-cut. If $\mathfrak{X}$ divides $X = \phi_x$, we write $X \equiv 0 \pmod{\mathfrak{X}}$ even when the X are not imbedded in a ring.

Ordinary Greek letters ϕ, ρ, μ, τ stand for functions.

4. Following Ore,[10] we define the structure Σ by means of its division relation $\supset$ and not by postulates on the cross-cut and union.

(4.1) $\mathfrak{X} \supset \mathfrak{X}$; if $\mathfrak{X} \supset \mathfrak{Y} \supset \mathfrak{Z}$ then $\mathfrak{X} \supset \mathfrak{Z}$. $\mathfrak{X} = \mathfrak{Y}$ if and only if $\mathfrak{X} \supset \mathfrak{Y}$ and $\mathfrak{Y} \supset \mathfrak{X}$. We write $\mathfrak{X} \supset \mathfrak{Y}$ or $\mathfrak{Y} \subseteq \mathfrak{X}$ if $\mathfrak{X} \supset \mathfrak{Y}$ or $\mathfrak{X} = \mathfrak{Y}$.

The cross-cut $[\mathfrak{X}, \mathfrak{Y}]$ is defined by

(4.2) $\mathfrak{X} \supseteq [\mathfrak{X}, \mathfrak{Y}]$, $\mathfrak{Y} \supseteq [\mathfrak{X}, \mathfrak{Y}]$; if $\mathfrak{X} \supseteq \mathfrak{Z}$ and $\mathfrak{Y} \supseteq \mathfrak{Z}$, then $[\mathfrak{X}, \mathfrak{Y}] \supseteq \mathfrak{Z}$.

The union is defined by

(4.3) $(\mathfrak{X}, \mathfrak{Y}) \supseteq \mathfrak{X}$, $(\mathfrak{X}, \mathfrak{Y}) \supseteq \mathfrak{Y}$; if $\mathfrak{Z} \supseteq \mathfrak{X}$ and $\mathfrak{Z} \supseteq \mathfrak{Y}$ then $\mathfrak{Z} \supseteq (\mathfrak{X}, \mathfrak{Y})$.

Let Ω be any sub-set of Σ. Since the division relation in Σ determines that in Ω, a pair of elements $\mathfrak{A}$, $\mathfrak{B}$ of Ω may have a cross-cut and union satisfying definitions (4.2), (4.3) for all elements of Ω. We write then $[\mathfrak{A}, \mathfrak{B}]_\Omega$, $(\mathfrak{A}, \mathfrak{B})_\Omega$ denoting the set relative to which we consider the cross-cut or union by a sub-script. Obviously

$$(\mathfrak{A}, \mathfrak{B}) \supseteq (\mathfrak{A}, \mathfrak{B})_\Omega \supseteq [\mathfrak{A}, \mathfrak{B}] \supseteq [\mathfrak{A}, \mathfrak{B}]_\Omega. \tag{4.4}$$

We call a set closed in this sense with respect to cross-cut and union *a structure within* Σ. If $(\mathfrak{A}, \mathfrak{B})_\Omega = (\mathfrak{A}, \mathfrak{B})$, $[\mathfrak{A}, \mathfrak{B}]_\Omega = [\mathfrak{A}, \mathfrak{B}]$ for every pair of elements of Ω, we call Ω *a sub-structure of* Σ. (Ore[10] p. 409.)

If any set of elements of Σ, finite or infinite have a unique cross-cut and union, we shall say that Σ is a *completely closed structure*.

II. Divisors of Arithmetical Functions and Their Ranks of Apparition

5. Let ϕ be an arithmetical function on $\mathfrak{o}$ to Σ. An element of Σ dividing one or more values of ϕ is said to divide ϕ or to be a *divisor of* ϕ. Obviously any factor of a divisor of ϕ divides ϕ. Hence:

---

[5] An arithmetic function need not however determine any structure homomorphism.

[6] The meaning of the symbols (...) and [...] is reversed in Ore's paper[10]. See section 4.

THEOREM 5.1. *The set of all divisors of an arithmetical function is closed under union.*

We denote this set by Δ.

Let $\mathfrak{A}$ be any divisor of ϕ. If $\mathfrak{A} \supseteq \phi_a$, we call a a *place of apparition of* $\mathfrak{A}$ *in* ϕ. We now assume

AXIOM 1.[7] *The places of apparition of every divisor of ϕ form an ideal of $\mathfrak{o}$.*

The ideal $\mathfrak{r}$ corresponding to a divisor $\mathfrak{Y}$ is called the *rank of apparition of* $\mathfrak{Y}$ *in* ϕ. We write $\mathfrak{r} = \rho(\mathfrak{Y})$. The set of all ranks of apparition is denoted by δ. The following theorem is now obvious.

THEOREM 5.2. *The correspondence between divisors and ranks of apparition defines an arithmetical function ρ on Δ to $\bar{\sigma}$.*

We cannot prove that the set Δ is closed under cross-cut as well as union. We assume:

AXIOM 2. *The set Δ of all divisors of ϕ is a completely closed sub-structure of Σ.*

THEOREM 5.3. *If $\mathfrak{A}$, $\mathfrak{B}$ are divisors of ϕ and $\mathfrak{M} = [\mathfrak{A}, \mathfrak{B}]$, then $[\rho(\mathfrak{A}), \rho(\mathfrak{B})] = \rho(\mathfrak{M})$.*

PROOF. By axiom 2, $\mathfrak{M}$ is a divisor of ϕ and by (4.2) $\mathfrak{A} \supseteq \mathfrak{M}$. Hence $\rho(\mathfrak{A}) \supseteq \rho(\mathfrak{M})$ by theorem 5.2. Similarly $\rho(\mathfrak{B}) \supseteq \rho(\mathfrak{M})$ so that $[\rho(\mathfrak{A}), \rho(\mathfrak{B})] \supseteq \rho(\mathfrak{M})$ by (4.2).

Assume that m lies in $[\rho(\mathfrak{A}), \rho(\mathfrak{B})]$. Then $m \equiv 0 \pmod{\rho(\mathfrak{A})}$, $m \equiv 0 \pmod{\rho(\mathfrak{B})}$ by (4.2). Hence $\phi_m \equiv 0 \pmod{\mathfrak{A}}$, $\phi_m \equiv 0 \pmod{\mathfrak{B}}$ or by (4.2) again, $\phi_m \equiv 0 \pmod{\mathfrak{M}}$. Hence $m \equiv 0 \pmod{\rho(\mathfrak{M})}$ so that $\rho(\mathfrak{M}) \supseteq [\rho(\mathfrak{A}), \rho(\mathfrak{B})]$.

THEOREM 5.31. "DECOMPOSITION THEOREM." *If the values of ϕ lie in a commutative ring $\mathfrak{O}$ with a unit element and if $\mathfrak{A}$, $\mathfrak{B}$ are any two divisors of ϕ such that $(\mathfrak{A}, \mathfrak{B}) = \mathfrak{O}$, then $\rho(\mathfrak{A}\mathfrak{B}) = [\rho(\mathfrak{A}), \rho(\mathfrak{B})]$.*

PROOF. If $(\mathfrak{A}, \mathfrak{B}) = \mathfrak{O}$, then $[\mathfrak{A}, \mathfrak{B}] = \mathfrak{A}\mathfrak{B}$ (van der Waerden,[11] p. 45).

If $\mathfrak{o}$ is a principal ideal ring, the following theorem allows us to replace axiom 1 by a simple structure condition. The result is independent of axiom 2.

THEOREM 5.4.[8] *If $\mathfrak{o}$ is a principal ideal ring, the places of apparition of every divisor of ϕ form an ideal if and only if ϕ defines a homomorphism with respect to union between $\mathfrak{o}$ and the values of ϕ in Σ.*

PROOF. If $\mathfrak{o}$ is a principal ideal ring, the union (m, n) of (m) and (n) is a principal ideal (d): we write $d = (m, n)$. Let $\mathfrak{A}$ be a divisor of ϕ, and assume that $\phi_m \equiv 0 \pmod{\mathfrak{A}}$, $\phi_n \equiv 0 \pmod{\mathfrak{A}}$. Also assume that $(m, n) = d$ in $\mathfrak{o}$ implies that $(\phi_m, \phi_n) = \phi_d$ in Σ. Then by (4.3) $\phi_d \equiv 0 \pmod{\mathfrak{A}}$. Now $d \mid m$, $d \mid n$. Hence $d \mid m - n$ so that $\phi_d \supseteq \phi_{m-n}$. Hence by (4.1) $\phi_{m-n} \equiv 0 \pmod{\mathfrak{A}}$. Thus if m and n are places of apparition of $\mathfrak{A}$, so is $m - n$. But since ϕ is arithmetical, if a is a place of apparition so is xa, x any element of $\mathfrak{o}$. Hence the places of apparition of $\mathfrak{A}$ form an ideal.

Assume conversely that $\mathfrak{o}$ is a principal ideal ring and that the places of

---

[7] It suffices to assume the places of apparition form a module, since ϕ is arithmetical.

[8] Given by Ward[12] for the case $\mathfrak{o}$ and Σ the ring of rational integers.

apparition of any divisor $\mathfrak{A}$ of ϕ form an ideal $\mathfrak{a} = (a)$. Let $(\phi_m, \phi_n) = D$, and let (d') be the rank of apparition of D in ϕ (Theorem 5.1). Since m, n lie in (d'), $d' \mid m$, $d' \mid n$. Hence if $(m, n) = d$, $d' \mid d$ by (4.3). Hence $\phi_d \equiv 0 \pmod{D}$. But since $d \mid m$, $d \mid n$, $\phi_d \supseteq \phi_m$, $\phi_d \supseteq \phi_n$ so that $\phi_d \supseteq D$ by (4.3). Thus $\phi_d = D$.

6. It is possible to have divisors $\mathfrak{A}$, $\mathfrak{B}$ of ϕ for which $\rho(\mathfrak{A}) = \rho(\mathfrak{B})$ $\mathfrak{A} \neq \mathfrak{B}$. Let $\mathfrak{M} = [\mathfrak{A}, \mathfrak{B}]$. Then by Theorem 5.3, $\rho(\mathfrak{M}) = [\rho(\mathfrak{A}), \rho(\mathfrak{B})] = \rho(\mathfrak{A})$. Thus the set of all elements $\mathfrak{Z}$ of Δ such that $\rho(\mathfrak{Z}) = n$ is closed under cross-cut. Hence by axiom 2, for every rank of apparition $\mathfrak{n}$ there exists a divisor $\mathfrak{N}$ of ϕ such that

$$\rho(\mathfrak{N}) = \mathfrak{n}; \text{ if } \mathfrak{C} \text{ divides } \phi \text{ and } \rho(\mathfrak{C}) = \mathfrak{n} \text{ then } \mathfrak{C} \supseteq \mathfrak{N}. \tag{6.1}$$

We call such a divisor a *maximal* divisor of ϕ.[9] We denote the set of all maximal divisors of ϕ by $\bar{\mu}$.

THEOREM 6.1. *Let $\mathfrak{A}$ be any divisor of ϕ and $\mathfrak{a}$ its rank of apparition. If $\mathfrak{N}$ is a maximal divisor of ϕ such that $\mathfrak{a} \supseteq \rho(\mathfrak{N})$, then $\mathfrak{A} \supseteq \mathfrak{N}$.*

PROOF. Let $[\mathfrak{A}, \mathfrak{N}] = \mathfrak{B}$. $\mathfrak{B}$ is a divisor of ϕ by axiom 2. Hence $\rho(\mathfrak{B})$ exists, and by theorem 5.3, $\rho(\mathfrak{B}) = [\rho(\mathfrak{A}), \rho(\mathfrak{N})] = \rho(\mathfrak{N})$ since $\rho(\mathfrak{A}) \supseteq \rho(\mathfrak{N})$. Since $\mathfrak{N}$ is maximal, $\mathfrak{B} \supseteq \mathfrak{N}$. But $\mathfrak{A} \supseteq \mathfrak{B}$.

As a corollary, we have

THEOREM 6.11. *If $\mathfrak{A}$ and $\mathfrak{B}$ are maximal divisors of ϕ with ranks of apparition $\mathfrak{a}$ and $\mathfrak{b}$ respectively, then $\mathfrak{A} \supseteq \mathfrak{B}$ when and only when $\mathfrak{a} \supseteq \mathfrak{b}$.*

THEOREM 6.2. *The set $\bar{\mu}$ of all maximal divisors of ϕ is closed under union.*

PROOF. Let $\mathfrak{D} = (\mathfrak{A}, \mathfrak{B})$. By (4.3),

(i) $\qquad \mathfrak{D} \supseteq \mathfrak{A}, \mathfrak{D} \supseteq \mathfrak{B}$;

(ii) $\qquad$ If $\mathfrak{C} \supseteq \mathfrak{A}$, $\mathfrak{C} \supseteq \mathfrak{B}$ then $\mathfrak{C} \supseteq \mathfrak{D}$.

By theorem 5.1, $\mathfrak{D}$ divides ϕ. Let $\mathfrak{d}$ be its rank of apparition and $\mathfrak{C}$ any divisor of ϕ such that $\rho(\mathfrak{C}) = \mathfrak{d}$. By (i) and theorem 5.2, $\rho(\mathfrak{C}) \supseteq \rho(\mathfrak{A})$. Hence since $\mathfrak{A}$ is maximal, $\mathfrak{C} \supseteq \mathfrak{A}$. Similarly $\mathfrak{C} \supseteq \mathfrak{B}$. Hence $\mathfrak{C} \supseteq \mathfrak{D}$ by (ii) so that $\mathfrak{D}$ is maximal.

THEOREM 6.21. *If $\mathfrak{A}$ and $\mathfrak{B}$ are maximal divisors of ϕ and $(\mathfrak{A}, \mathfrak{B}) = \mathfrak{D}$ then $(\rho(\mathfrak{A}), \rho(\mathfrak{B}))_\delta$ exists and equals $\rho(\mathfrak{D})$.*

PROOF. By theorem 6.2, $\mathfrak{D}$ is maximal. Hence by theorem 6.11, if $\mathfrak{a} = \rho(\mathfrak{A})$, $\mathfrak{b} = \rho(\mathfrak{B})$, $\mathfrak{c} = \rho(\mathfrak{C})$ and $\mathfrak{d} = \rho(\mathfrak{D})$, (i) $\mathfrak{d} \supseteq \mathfrak{a}$, $\mathfrak{d} \supseteq \mathfrak{b}$; (ii) if $\mathfrak{c} \supseteq \mathfrak{a}$, $\mathfrak{c} \supseteq \mathfrak{b}$ then $\mathfrak{c} \supseteq \mathfrak{d}$. (i) and (ii) are the definition of union.

Since to every rank of apparition corresponds a maximal divisor, in view of theorem 6.21 and 5.3, we can state

THEOREM 6.211. *The ranks of apparition of the divisors of ϕ form a structure within $\bar{\sigma}$.*

[9] For example let ϕ_n denote the n^{th} of the Fibonacci numbers 1, 1, 2, 3, 5, 8, $\cdots$. Then $\phi_{19} = 4181 = 37 \times 113$. Hence for $\mathfrak{A} = (37)$, $\mathfrak{B} = (113)$, $\mathfrak{n} = \rho(\mathfrak{A}) = \rho(\mathfrak{B}) = (19)$ while $\mathfrak{N} = (4181)$.

THEOREM 6.3. *If $\mathfrak{A}$ and $\mathfrak{B}$ are maximal divisors of ϕ, then the cross-cut $\mathfrak{M} = [\mathfrak{A}, \mathfrak{B}]_\mu$ within $\tilde{\mu}$ exists, and $\rho(\mathfrak{M}) = [\rho(\mathfrak{A}), \rho(\mathfrak{B})]$.*

PROOF. Let $[\mathfrak{A}, \mathfrak{B}] = \mathfrak{N}$. By axiom 2, $\mathfrak{N}$ divides ϕ. Let $\mathfrak{m} = \rho(\mathfrak{N})$ and let $\mathfrak{M}$ be the corresponding maximal divisor such that $\mathfrak{m} = \rho(\mathfrak{M})$. Then $\mathfrak{N} \supseteq \mathfrak{M}$. I say that $\mathfrak{M} = [\mathfrak{A}, \mathfrak{B}]_\mu$. For $\mathfrak{A} \supseteq \mathfrak{M}$, $\mathfrak{B} \supseteq \mathfrak{M}$ by (4.4), (4.2). Let $\mathfrak{C}$ be any other maximal divisor such that $\mathfrak{A} \supseteq \mathfrak{C}$, $\mathfrak{B} \supseteq \mathfrak{C}$. Then $\mathfrak{N} \supseteq \mathfrak{C}$ by (4.2). Hence $\rho(\mathfrak{N}) \supseteq \rho(\mathfrak{C})$ by theorem 6.2, so that $\rho(\mathfrak{M}) \supseteq \rho(\mathfrak{C})$. Since $\mathfrak{M}$ and $\mathfrak{C}$ are maximal, $\mathfrak{M} \supseteq \mathfrak{C}$ by theorem 6.1, so that $\mathfrak{M} = [\mathfrak{A}, \mathfrak{B}]_\mu$ by (4.2).

THEOREM 6.31. *The maximal divisors of ϕ form a structure within Σ isomorphic* (Ore,[10] pp. 416–418) *to the structure of ranks of apparition within $\tilde{\sigma}$.*

PROOF. Theorems 6.3, 6.2, 6.11.

III. MODULAR PERIODICITY

7. We assume henceforth that the values of ϕ lie in a commutative ring $\mathfrak{O}$. Σ now denotes the structure of all ideals of $\mathfrak{O}$, so that the values of ϕ *quâ* elements of Σ are principal ideals of $\mathfrak{O}$.

For the present ϕ is any function on $\mathfrak{o}$ to $\mathfrak{O}$, not necessarily arithmetical, and $\tilde{\sigma}$ denotes the structure of all *modules* of $\mathfrak{o}$. We use small German letters now for modules instead of ideals. If $\mathfrak{m}$ is a module, $m \equiv 0 \pmod{\mathfrak{m}}$ means $\mathfrak{m}$ contains m.

Let $\mathfrak{S}$ be any element of Σ. If there exists an element $m \neq 0$ of $\mathfrak{o}$ such that

$$\phi_{x+m} \equiv \phi_x \pmod{\mathfrak{S}}, \qquad \text{every } x \text{ of } \mathfrak{o},^{10} \tag{7.1}$$

$\mathfrak{S}$ is called a *modulus* of ϕ, and m a *period* of ϕ modulo $\mathfrak{S}$. The periods (0 included) obviously form a module, $\mathfrak{s}$ *the characteristic module of ϕ modulo $\mathfrak{S}$*. We shall also call $\mathfrak{s}$ the module of $\mathfrak{S}$, writing

$$\mathfrak{s} = \mu(\mathfrak{S}). \tag{7.2}$$

Let Π denote the set of all moduli of ϕ in Σ, and $\tilde{\pi}$ the set of all characteristic modules in $\tilde{\sigma}$. Clearly as in section 5, we have

THEOREM 7.1. *Π is closed under union.*

THEOREM 7.2. *μ is an arithmetic function on Π to $\tilde{\sigma}$.*

We cannot prove in general that Π is closed under cross-cut. We assume:

AXIOM 3. *The set of all moduli of ϕ is a completely closed sub-structure of Σ.*

Then the following theorems follow precisely as in section 5.

THEOREM 7.3. *If $\mathfrak{A}$, $\mathfrak{B}$ are moduli of ϕ and if $\mathfrak{M} = [\mathfrak{A}, \mathfrak{B}]$, then $\mu(\mathfrak{M}) = [\mu(\mathfrak{A}), \mu(\mathfrak{B})]$.*

THEOREM 7.31. "DECOMPOSITION THEOREM."[11] *If $\mathfrak{A}$, $\mathfrak{B}$ are moduli of ϕ and $(\mathfrak{A}, \mathfrak{B}) = \mathfrak{O}$, then $\mu(\mathfrak{A}\mathfrak{B}) = [\mu(\mathfrak{A}), \mu(\mathfrak{B})]$.*

---

[10] We shall omit this phrase henceforth, reserving the letter x exclusively to denote every element of $\mathfrak{o}$.

[11] Stated in Ward[13] for linear recurring series. The analogous theorems 5.31 and 10.31 appear to be new.

8. The theory of maximal moduli exactly parallels the theory of maximal divisors.

For every possible module of periods $\mathfrak{n}$ there exists a maximal modulus $\mathfrak{N}$ such that

$$(8.1) \qquad \mu(\mathfrak{N}) = \mathfrak{n}; \qquad \text{if} \qquad \mu(\mathfrak{S}) = \mathfrak{n},\ \mathfrak{S} \text{ a module, } \mathfrak{S} \supseteq \mathfrak{N}.$$

THEOREM 8.1. *Let $\mathfrak{A}$ be any modulus of ϕ, $\mathfrak{a}$ its characteristic module. Then if $\mathfrak{N}$ is a maximal modulus such that $\phi \supseteq \mathfrak{a}(\mathfrak{N})$, $\mathfrak{A} \supseteq \mathfrak{N}$.*

THEOREM 8.11. *If $\mathfrak{A}$ and $\mathfrak{B}$ are maximal moduli of ϕ with characteristic modules $\mathfrak{a}$ and $\mathfrak{b}$ respectively, then $\mathfrak{A} \supseteq \mathfrak{B}$ when and only when $\mathfrak{a} \supseteq \mathfrak{b}$.*

THEOREM 8.2. *The set of all maximal moduli of ϕ is closed under union.*

THEOREM 8.21. *If $\mathfrak{A}$ and $\mathfrak{B}$ are maximal moduli of ϕ and $(\mathfrak{A}, \mathfrak{B}) = \mathfrak{D}$, then $(\mu(\mathfrak{A}), \mu(\mathfrak{B}))_\tau$ exists and equals $\mu(\mathfrak{D})$.*

THEOREM 8.211. *The characteristic modules of the moduli of ϕ form a structure within $\bar{\sigma}$.*

THEOREM 8.3. *If $\mathfrak{A}$ and $\mathfrak{B}$ are maximal moduli of ϕ then $\mathfrak{M} = [\mathfrak{A}, \mathfrak{B}]_\tau$ exists, and $\mu(\mathfrak{M}) = [\mu(\mathfrak{A}), \mu(\mathfrak{B})]$.*

THEOREM 8.31. *The maximal moduli of ϕ form a structure within Σ isomorphic to the structure of characteristic modules within $\bar{\sigma}$.*

It is easily seen that if we assume in analogy with axiom 1,

AXIOM 4. *The periods of any modulus of ϕ form an ideal.*

Then all the theorems of this section and the preceding one hold if the word module is everywhere replaced by ideal, and the notation $\mathfrak{a}, \cdots, \mathfrak{z}$ is understood to mean ideals instead of modules.

IV. RESTRICTED PERIODS. RELATIONSHIPS BETWEEN DIVISORS AND MODULI OF AN ARITHMETIC FUNCTION

9. The concept of "restricted period" which we formulate abstractly here is very important in the theory of the Lucas function and linear recurring series in general. (Carmichael,[2] Ward[4].) As in sections 7, 8 we assume that the values of ϕ lie in a commutative ring $\mathfrak{O}$ containing a unit element. Let $\mathfrak{S}$ be an ideal of $\mathfrak{O}$. If there exist elements M and m of $\mathfrak{O}$ and $\mathfrak{o}$ such that

$$(9.1) \qquad \phi_{x+m} \equiv M\phi_x \pmod{\mathfrak{S}} \text{ all } x \text{ in } \mathfrak{o}$$

$$(9.11) \qquad ((M), \mathfrak{S}) = \mathfrak{O}$$

then m is called a *restricted period* of ϕ modulo $\mathfrak{S}$, and M a *multiplier* of ϕ modulo $\mathfrak{S}$.

THEOREM 9.1. *The multipliers of ϕ modulo $\mathfrak{S}$ are closed under multiplication. The restricted periods of ϕ modulo $\mathfrak{S}$ form an additive semi-group.*

PROOF. Assume that $\phi_{x+m_i} \equiv M_i\phi_x \pmod{\mathfrak{S}}$, $((M_i), \mathfrak{S}) = \mathfrak{O}$, $i = 1, 2$. Then $\phi_{x+m_1+m_2} \equiv M_1\phi_{x+m_2} \equiv M_1M_2\phi_x \pmod{\mathfrak{S}}$. Also (Van der Waerden,[11] p. 45) $((M_1M_2), \mathfrak{S}) = \mathfrak{O}$.

We shall now assume

AXIOM 5.[12] *For any modulus $\mathfrak{S}$ the multipliers of ϕ modulo $\mathfrak{S}$ form a multiplicative group in $\mathfrak{O}$* (Ward,[13] p. 162).

We denote this group by $\mathfrak{G}$.

THEOREM 9.11. *The restricted periods of ϕ modulo $\mathfrak{S}$ form a module.*

PROOF. With the notation of theorem 9.1, let N_1 be the inverse of M_1 in $\mathfrak{G}$, so that $N_1M_1 \equiv 1 \pmod{\mathfrak{S}}$. Then by (7.2)

$$\phi_{x+m_2-m_1} \equiv N_1M_1\phi_{x-m_1+m_2} \equiv N_1M_1M_2\phi_{x-m_1} \equiv N_1M_2\phi_x \pmod{\mathfrak{S}}.$$

By axiom 5, theorem 7.1 N_1M_2 is a multiplier so that $m_2 - m_1$ is a period.

THEOREM 9.2. *Let $\mathfrak{S}$ be a modulus, $\mathfrak{G}$ its group, $\mathfrak{s}$ its restricted period. Let $\mathfrak{T}$ be any divisor of $\mathfrak{S}$. Then $\mathfrak{T}$ also is a modulus. If $\mathcal{H}$ is its group and* t *its restricted period,* $\mathfrak{G} \subseteq \mathcal{H}$, $\mathfrak{t} \supseteq \mathfrak{s}$.

PROOF. Clear.

THEOREM 9.3. *The set of all moduli of ϕ and the set of all moduli of restricted periods are identical.*

PROOF. If $\mathfrak{S}$ is an ordinary modulus, its group of multipliers consists of the single element 1 of $\mathfrak{O}$. On the other hand, if $\mathfrak{S}$ is a modulus of restricted periods, $\mathfrak{S}$ has the multiplier 1 by axiom 5 and hence is an ordinary modulus.

THEOREM 9.31.[13] *If $\tau(\mathfrak{S})$ and $\mu(\mathfrak{S})$ are respectively the module of restricted periods and the module of periods of the modulus $\mathfrak{S}$ of ϕ, then* $\tau(\mathfrak{S}) \supseteq \mu(\mathfrak{S})$.

PROOF. Clear from Theorem 9.3 and axiom 5.

10. In view of theorem 9.3, it is unnecessary to show that the moduli of the restricted periods of ϕ are closed under union, and theorem 9.2 makes it evident that τ is an arithmetical function on $\bar{\pi}$ to $\bar{\sigma}$. Even if we retain axiom 3 of section 7, we cannot prove the analogue of theorems 5.3 and 7.3 viz.:—"If $\mathfrak{A}$, $\mathfrak{B}$ are moduli of ϕ and if $\mathfrak{M} = [\mathfrak{A}, \mathfrak{B}]$, then $\tau(\mathfrak{M}) = [\tau(\mathfrak{A}), \tau(\mathfrak{B})]$." We must content ourselves with $\tau(\mathfrak{M}) \subseteq [\tau(\mathfrak{A}), \tau(\mathfrak{B})]$.

To see why this lack occurs, let us try to show that $[\tau(\mathfrak{A}), \tau(\mathfrak{B})] \subseteq \tau(\mathfrak{M})$. If m is any element of $[\tau(\mathfrak{A}), \tau(\mathfrak{B})]$, we infer from (4.2) that

$$\phi_{x+m} \equiv M\phi_x \pmod{\mathfrak{A}}, \qquad \phi_{x+m} \equiv N\phi_x \pmod{\mathfrak{B}}$$

$$((M), \mathfrak{A}) = \mathfrak{O}, \qquad ((N), \mathfrak{B}) = \mathfrak{O}.$$

To show that m lies in $\tau(\mathfrak{M})$, it is necessary and sufficient to show that there exists an element S of $\mathfrak{O}$ such that

$$\phi_{x+m} \equiv S\phi_x \pmod{\mathfrak{M}}, \qquad ((S), \mathfrak{M}) = \mathfrak{O}.$$

Since $\mathfrak{A} \supseteq \mathfrak{M}$, $\mathfrak{B} \supseteq \mathfrak{M}$, we must also have

$$S \equiv M \pmod{\mathfrak{A}}, \qquad S \equiv N \pmod{\mathfrak{B}}.$$

---

[12] This axiom always holds if $\mathfrak{O}$ is a ring of algebraic integers or more generally whenever the ring $\mathfrak{O}/\mathfrak{S}$ is finite.

[13] Generalizes Carmichael[2], p. 355.

Now such an element S need not exist. For example, take for $\mathfrak{O}$ the ring of rational integers, and let $\mathfrak{A} = (6)$, $\mathfrak{B} = (9)$, $M = 5$, $N = 4$. Then $\mathfrak{M} = (18)$, and no S exists.

On the other hand if $(\mathfrak{A}, \mathfrak{B}) = \mathfrak{O}$, such an S does exist since the Chinese remainder theorem holds in a commutative ring with unit element (van der Waerden,[11] p. 85). Thus a "decomposition theorem" holds for the restricted period analogous to theorems 5.31 and 7.31.

Decomposition Theorem 10.31. *If* $\mathfrak{A}$, $\mathfrak{B}$ *are moduli and* $(\mathfrak{A}, \mathfrak{B}) = \mathfrak{O}$, *then* $\tau(\mathfrak{A}\mathfrak{B}) = [\tau(\mathfrak{A}), \tau(\mathfrak{B})]$.

Since theorem 7.3 fails, we cannot introduce maximal restricted period moduli, and the theorems of sections 6 and 8 have no analogues.

11. It remains to discuss the relationship between the rank of apparition and restricted period of any element of Σ. Let us suppose that ϕ is an arithmetic function on $\mathfrak{o}$ to $\mathfrak{O}$, and that axioms 1, 2, 3 and 5 hold.

Consider the elements ϕ_0 and ϕ_1. Since $x \mid 0$, $1 \mid x$ for every x of $\mathfrak{o}$, $\phi_x \supseteq \phi_0$ and $\phi_1 \supseteq \phi_x$ for every value of ϕ in $\mathfrak{O}$. The simplest way to satisfy these two conditions is for ϕ_0 to equal the zero element and ϕ_1 the unit element of the ring $\mathfrak{O}$. An arithmetic function with this property will be called *normal*. We assume

Axiom 6. *ϕ is a normal arithmetic function on $\mathfrak{o}$ to $\mathfrak{O}$.*

Theorem 11.1.[14] *Every modulus of ϕ is a divisor of ϕ, and the rank of apparition of any modulus divides its restricted period.*

Proof. Let $\mathfrak{S}$ be any modulus, m a restricted period of $\mathfrak{S}$. Then since

$$\phi_{m+x} \equiv M\phi_x \pmod{\mathfrak{S}}, \qquad \phi_m \equiv M\phi_0 \equiv 0 \pmod{\mathfrak{S}}.$$

Hence $m \equiv 0 \bmod \rho(\mathfrak{S})$.

California Institute of Technology.

REFERENCES

[1] E. Lucas, *Amer. Jour. of Math.* vol. 1 (1878) pp. 184–239, 289–321.
[2] R. D. Carmichael, *Quarterly Journal* vol. 48 (1920) pp. 343–372.
[3] H. T. Engstrom, *Trans. Amer. Math. Soc.* vol. 33 (1931) pp. 210–218.
[4] M. Ward, *Trans. Amer. Math. Soc.* vol. 35 (1933) pp. 600–628.
[5] D. H. Lehmer, *Annals of Math.* (2), vol. 31 (1930) pp. 419–448.
[6] M. Hall, *Amer. Jour. of Math.* vol. 58 (1936) pp. 577–584.
[7] M. Ward, *Trans. Amer. Math. Soc.* (reference later).
[8] L. E. Dickson, *History*, vol. 1, chapter V.
[9] G. H. Halphen, *Traite des Fonctions Elliptiques*, Part I, chap. IV.
[10] O. Ore, *Annals of Math.* (2) vol. 36 (1935), pp. 406–437.
[11] Van der Waerden, *Modern Algebra Part* 2, Berlin (1931).
[12] M. Ward, *Bulletin Amer. Math. Soc.* (reference later).
[13] M. Ward, *Trans. Amer. Math. Soc.* vol. 33 (1931) pp. 153–165.

[14] This result generalizes a theorem of Hall[6] on linear divisibility sequences. For the Lucas function, the rank of apparition and restricted period are equal.

Annals of Mathematics
Vol. 38, No. 3, July, 1937

SUR L'ADDITION HOMOLOGIQUE DES TYPES DE TRANSFORMATIONS CONTINUES EN SURFACES SPHÉRIQUES

Par Karol Borsuk

(Received February 1, 1937)

Deux transformations continues φ et ψ d'un espace métrique séparable M en une surface sphérique n-dimensionnelle S_n sont dites *homologues*, lorsque tout vrai cycle $\mathfrak{C}$ de M[1] est transformé par φ et ψ en vrais cycles $\mathfrak{C}_\varphi$ et $\mathfrak{C}_\psi$ homologues dans S_n. L'ensemble S_n^M[2] de toutes les transformations continues de M en S_n se décompose ainsi en classes des transformations homologues dites *types d'homologie*[3] ou *h-types*. Pour tout $\varphi \in S_n^M$ le h-type contenant φ sera designé par $[\varphi]$. Une transformation $\chi \in S_n^M$ sera dite *h-somme* des transformations $\varphi, \psi \in S_n^M$, lorsque pour tout vrai cycle $\mathfrak{C}$ de M on a $\mathfrak{C}_\chi \sim \mathfrak{C}_\varphi + \mathfrak{C}_\psi$ dans S_n.[4] Il résulte de cette définition que le h-type de la h-somme χ de φ et ψ est défini d'une manière univoque par les h-types $[\varphi]$ et $[\psi]$. Ceci nous permet de traiter le h-type $[\chi]$ comme la *somme homologique* (*h-somme*) des types $[\varphi]$ et $[\psi]$.

L'univocité de cette somme étant une conséquence immédiate de sa définition, il n'en est pas ainsi quant à son existence. Toutes les fonctions $\varphi \in S_n^M$ étant dans le cas $\dim M < n$ homologues à une constante, la question d'existence de la h-somme ne se pose que pour les transformations en S_n des espaces de

[1] Un complexe algébrique (aux coefficients arbitraires) K est dit *ϵ-complexe de M*, lorsque tous ses sommets appartiennent à M et que la distance maximum entre deux sommets d'un simplexe de K est $< \epsilon$. Deux ϵ-complexes K_1 et K_2 de M sont dits *η-homologues dans M* (notation: $K_1 \sim_\eta K_2$ dans M), lorsqu'il existe un η complexe K de M dont la frontière est égale à $K_1 - K_2$. Une suite $\mathfrak{C} = \{C_i\}$ s'appelle *vrai cycle de M*, lorsqu'il existe un sous-ensemble compact M_0 de M tel que C_i est un ϵ_i-cycle de M_0, avec $\lim_{i \to \infty} \epsilon_i = 0$, pour tout $i = 1, 2, \cdots$. Dans le cas où la dimension de C_i est 0, nous admettons en outre que la somme de ses coefficients est égale à 0 (c.à d.nous ne considérons que des cycles qui sont "*berandungsfähig*"). Par la somme de deux vrais cycles $\mathfrak{C} = \{C_i\}$ et $\mathfrak{C}' = \{C_i'\}$ de M on entend le vrai cycle $\mathfrak{C} + \mathfrak{C}' = \{C_i + C_i'\}$. Deux cycles $\mathfrak{C}$ et $\mathfrak{C}'$ de M sont dits *homologues dans M* (notation: $\mathfrak{C} \sim \mathfrak{C}'$ dans M) lorsqu'il existe un sous-ensemble compact M_0 de M et une suite $\{\eta_i\}$ de nombres positifs convergente vers 0, pour lesquels $C_i \sim_{\eta_i} C_i'$ dans M_0 pour tout $i = 1, 2, \cdots$. Chaque fonction continue φ transformant M en un autre espace N fait correspondre aux cycles C_i de M les cycles $C_{i,\varphi}$ de N qui constituent un vrai cycle $\{C_{i,\varphi}\}$ de N, désigné par $\mathfrak{C}_\varphi$.

[2] N^M désigne l'ensemble de toutes les transformations continues de M en sous-ensembles de N. Dans le cas où N est compact, on considère N^M comme un espace métrique en posant $\rho(\varphi, \psi) = \operatorname{Sup}_{x \in M} \rho\,[\varphi(x), \psi(x)]$ pour tout $\varphi, \psi \in N^M$.

[3] Comp. P. Alexandroff et H. Hopf, *Topologie* I, Berlin 1935, p. 321.

[4] Les définitions des fonctions homologues, du type d'homologie et de la h-somme sont valables sans aucune modification dans le cas des transformations de M en un espace arbitraire N.

dimension $\geqq n$. Les résultats récents de M. H. Freudenthal[5] donnent une réponse affirmative à cette question dans le cas ou l'espace M est compact et sa dimension est $\leqq n$. Ensuite j'ai donné[6] un procédé plus général qui permet en particulier de démontrer l'existence de la somme en question dans le cas où la dimension de l'espace M (compact ou non) est $< 2n - 1$.

Dans cet ordre d'idées je vais démontrer dans ce travail que la h-sommation des h-types de S_n^M se laisse toujours effectuer quand dim $M < 2n$. Il se trouve en outre que, grâce aux résultats récents de M. H. Hopf, cette dernière inégalité ne se laisse remplacer dans le cas général par aucune inégalité moins restrictive.

1. Lemme. *Soit $a \in S_n$ et $b, b' \in S_n - (a)$. L'ensemble $[S_n \times (a)] + [(a) \times S_n]$*[7] *est un rétracte par déformation*[8] *de l'ensemble $S_n \times S_n - (b, b')$.*

Démonstration. La surface sphérique S_n se laisse obtenir du cube euclidien n-dimensionnel $Q_n = \underset{(x_1 x_2 \cdots x_n)}{E} [\,|\,x_i\,| \leqq 1; i = 1, 2, \cdots n]$ par l'identification de tous les points de sa frontière. On peut admettre que c'est le point a qui dérive de cette identification. Les points b et b' appartiennent alors à l'intérieur de Q_n et par conséquent $p_0 = (b, b')$ est un point intérieur du cube euclidien $Q_{2n} = Q_n \times Q_n$. Il en résulte que la projection p^* de p du centre p_0 sur la frontière de Q_{2n} est définie d'une manière univoque pour tout $p \in Q_{2n} - (p_0)$. En posant

$g(p, t) =$ *point du segment $p_0 p^*$ tel que* $\rho(p, g(p, t)) = t \cdot \rho(p, p^*)$, on obtient une fonction rétractant par déformation l'ensemble $Q_{2n} - (p_0)$ en la frontière du cube Q_{2n}. Remarquons maintenant qu'on déduit l'ensemble $S_n \times S_n$ du cube Q_{2n} en identifiant des points (x, y) et (x', y') lorsque ils ne se distinguent que par les "coordonnées" appartenant à la frontière de Q_n. La frontière de Q_{2n} est transformée par cette identification en l'ensemble $[S_n \times (a)] + [(a) \times S_n]$ et par conséquent, la fonction $g(p, t)$ fait la rétraction demandée.

2. Théorème 1. *Si la dimension de l'espace M est $< 2n$, la h-somme existe pour chaque couple de h-types de S_n^M.*

---

[5] H. Freudenthal, *Die Hopfsche Gruppe*, Compositio Math. 2 (1935), p. 134–162.

[6] K. Borsuk, *Sur les groupes des classes de transformations continues*, C.R.de l'Ac. des Sc.202 (1936), p. 1401–1403. Les difficultés que l'on rencontre dans la définition de h-somme des transformations continues *univoques* n'existent pas pour les transformations continues *multivoques*. La somme homologique se laisse définir pour ces dernières transformations sans aucune hypothèse sur la dimension. Voir S. Lefschetz, *Sur les transformations des complexes en sphères*, Fund. Math. 27 (1936), p. 106.

[7] $M \times N$ désigne le *produit cartésien* des espaces M et N, c.à d.l'ensemble de tous les couples (x, y) où $x \in M$ et $y \in N$, métrisé par la formule

$$\rho[(x, y), (x', y')] = \sqrt{\rho(x, x')^2 + \rho(y, y')^2}.$$

[8] A est dit *rétracte de M par déformation*, lorsqu'il existe une fonction continue $g(x, t)$ de deux variables $x \in M$ et $0 \leqq t \leqq 1$ dont les valeurs appartiennent à M et telle que $g(x, 0) = x$ pour $x \in M$, $g(M, 1) = A$ et $g(x, 1) = x$ pour $x \in A$.

DÉMONSTRATION. Soient $\varphi, \psi \in S_n^M$. Il faut démontrer qu'il existe une fonction $\chi \in S_n^M$ constituant une h-somme pour les transformations φ et ψ. En considérant le couple $(\varphi(x), \psi(x))$ comme une seule transformation $f(x) = (\varphi(x), \psi(x))$ de M en $S_n \times S_n$ on obtient une correspondance biunivoque et bicontinue entre $S_n^M \times S_n^M$ et $(S_n \times S_n)^M$. L'homotopie[9] des fonctions $f(x) = (\varphi(x), \psi(x))$ et $f'(x) = (\varphi'(x), \psi'(x))$ dans $(S_n \times S_n)^M$ est par conséquent équivalente avec l'homotopie simultanée de φ avec φ' et de ψ avec ψ' dans S_n^M. La dimension de M étant $< 2n$, il existe[10] une fonction f'' homotope à f dans $(S_n \times S_n)^M$ et telle que la dimension de $f''(M)$ est $< 2n$. Il en résulte qu'il existe pour un point a arbitrairement donné dans S_n un point $(b, b') \in (S_n \times S_n) - [f''(M) + (a) \times S_n + S_n \times (a)]$. Il existe en outre, d'après le lemme, une déformation continue $g(p, t)$ de l'ensemble $S_n \times S_n - (b, b')$ en lui-même telle que $g(p, t)$ est une fonction rétractant $S_n \times S_n - (b, b')$ en $[S_n \times (a)] + [(a) \times S_n]$. La formule $f_t(x) = g[f''(x), t]$ fait correspondre à cette déformation une déformation continue de la fonction $f_0 = f''$ en une fonction f_1 dont les valeurs appartiennent à $[S_n \times (a)] + [(a) \times S_n]$. Or f_1 est de la forme (φ', ψ') où φ' et ψ' sont respectivement homotopes (et par conséquent aussi homologues) à φ et ψ et satisfont à la condition: pour tout $x \in M$ une des valeurs $\varphi'(x)$ et $\psi'(x)$ est égale à a. Par conséquent les ensembles fermés $M_1 = \underset{x}{E}[\varphi(x) = a]$ et $M_2 = \underset{x}{E}[\psi(x) = a]$ constituent une décomposition de M telle que

$$\varphi'(M_1) = \psi'(M_2) = a. \tag{1}$$

Il en résulte qu'on obtient une fonction $\chi \in S_n^M$ lorsqu'on pose

$$\begin{aligned} \chi(x) &= \psi'(x) \text{ pour } x \in M_1, \\ \chi(x) &= \varphi'(x) \text{ pour } x \in M_2. \end{aligned} \tag{2}$$

Les fonctions φ' et ψ' étant homologues respectivement à φ et ψ, il ne reste qu'à démontrer que la fonction χ définie dans $M = M_1 + M_2$ par les formules (2) constitue une h-somme de fonctions $\varphi', \psi' \in S_n^M$ satisfaisant à la condition (1).[11]

On prouve facilement[12] qu'il existe pour tout vrai cycle $\mathfrak{C} = \{C_i\}$ de M un

---

[9] Deux fonctions $\varphi, \psi \in N^M$ sont dites *homotopes dans* N^M, lorsqu'il existe une famille de fonctions $\{\varphi_t\} \subset N^M$ dépendant d'une manière continue du parametre $0 \leqq t \leqq 1$ et telle que $\varphi_0 = \varphi$ et $\varphi_1 = \psi$. Une fonction homotope à une constante est dite *homotope à* 0. On sait que deux fonctions homotopes sont aussi homologues.

[10] C'est une conséquence facile d'un théorème de M. W. Hurewicz, sur les transformations continues d'un espace métrique séparable M en S_n. Voir son travail: *Über Abbildungen topologischer Räume auf die n-dimensionale Sphäre*, Fund. Math. 24 (1935), p. 146. Cf. aussi mon travail: *Sur les transformations continues n'augmentant pas la dimension*, Fund. Math. 28 (1937), p. 91, th. 3.

[11] Cette partie de la démonstration reste valable pour les transformations continues aux valeurs arbitraires (et non seulement pour les transformations en surfaces sphériques).

[12] Voir par exemple mon travail: *O zagadnieniu topologicznego scharakteryzowania sfer euklidesowych* (en polonais), Wiadomości Matematyczne 38 (1934), p. 12.

vrai cycle homologue à $\mathfrak{C}$ dans M tel que tous les sommets de chacun de ses simplexes appartiennent à un des ensembles M_ν. Or, on peut admettre que c'est le vrai cycle $\mathfrak{C}$ qui satisfait à cette condition. On a par conséquent $C_i = K_i^{(1)} + K_i^{(2)}$, où $K_i^{(\nu)}$ est un complexe de M_ν. Les fonctions φ' et ψ' transforment $\mathfrak{C}$ en vrais cycles $\mathfrak{C}_{\varphi'} = \{C_{i,\varphi'}\}$ et $\mathfrak{C}_{\psi'} = \{C_{i,\psi'}\}$ de S_n, où $C_{i,\varphi'} = K_{i,\varphi'}^{(1)} + K_{i,\varphi'}^{(2)}$ et $C_{i,\psi'} = K_{i,\psi'}^{(1)} + K_{i,\psi'}^{(2)}$. Dans le cas où la dimension du cycle C_i est $\geqq 1$, tous les simplexes du complexe $K_i^{(1)}$ sont transformés par φ' en 0 et par conséquent $K_{i,\varphi'}^{(1)} = 0$; pareillement $K_{i,\psi'}^{(2)} = 0$. Il vient:

$$C_{i,\varphi'} + C_{i,\psi'} = K_{i,\psi'}^{(1)} + K_{i,\varphi'}^{(2)}. \tag{3}$$

Or, cette relation persiste aussi dans le cas où la dimension du cycle C_i est 0. En effet, en désignant par α_i la somme des coefficients du complexe $K_i^{(\nu)}$ on a $C_{i,\varphi'} = \alpha_1 a + K_{i,\varphi'}^{(2)}$ et $C_{i,\psi'} = K_{i,\psi'}^{(1)} + \alpha_2 a$, ce qui entraîne (3), car la somme $\alpha_1 + \alpha_2$, égale à la somme des coefficients du cycle C_i, disparaît.[1] Pour achever la démonstration il ne reste donc qu'à remarquer que le cycle $C_{i,\chi}$ en lequel la fonction χ transforme C_i est, d'après (2), égal à $K_{i,\psi'}^{(1)} + K_{i,\varphi'}^{(2)}$.

3. Remarquons maintenant que:

1° L'opération de h-sommation est associative et commutative,

2° Le h-type de la fonction constante φ_0 joue pour l'opération de h-sommation le rôle de l'élément 0, car φ_0 transforme chaque vrai cycle de M en 0.

3° α désignant une fonction continue transformant S_n en elle-même avec le grade -1, on a pour tout vrai cycle $\mathfrak{C}$ de S_n l'homologie dans S_n entre les vrais cycles $\mathfrak{C}_\alpha$ et $-\mathfrak{C}$. Par conséquent, en posant $\varphi'(x) = \alpha\varphi(x)$ pour tout $\varphi \in S_n^M$ on obtient une fonction opposée à φ, c.à d. une fonction telle que la fonction constante constitue une h-somme de φ et φ'.

Les propositions 1°, 2° et 3° et le théorème 1 nous conduisent au corollaire suivant:

Corollaire. *Les h-types de transformations d'un espace M de dimension $< 2n$ en S_n constituent un groupe abélien dans lequel la h-sommation joue le rôle de l'addition.*

4. Nous allons maintenant montrer que l'inégalité dim $M < 2n$ ne se laisse pas en général remplacer dans le théorème 1 par une inégalité moins restrictive. On a notamment le théorème suivant:

Théorème 2. *On peut trouver pour tout nombre pair n des transformations continues de la variété $2n$-dimensionnelle $M = S_n \times S_n$ en S_n pour lesquelles une h-somme n'existe pas.*

Posons

$$\varphi(x, y) = \psi(y, x) = x \text{ pour tout } (x, y) \in S_n \times S_n. \tag{4}$$

Soit a un point arbitrairement donné de S_n. La fonction φ transforme la surface sphérique $S_n \times (a)$ en S_n avec le grade 1 et la surface sphérique $(a) \times S_n$ avec le grade 0, tandis que ψ transforme $S_n \times (a)$ avec le grade 0 et $(a) \times S_n$

avec le grade 1. Or une h-somme $\chi_0 \in S_n^{S_n \times S_n}$ de φ et ψ transformerait chacune des surfaces sphériques $S_n \times (a)$ et $(a) \times S_n$ avec le grade 1. Mais dans le cas ou n est un nombre pair, une telle fonction χ_0 n'existe pas.[13]

5. Nous avons démontré dans le Numéro précédent que le manque d'une transformation continue de $S_n \times S_n$ en S_n transformant chacune des surfaces sphériques $S_n \times (a)$ et $(a) \times S_n$ en S_n avec le grade 1 entraîne qu'il n'existe pas de h-somme pour les fonctions φ et ψ définies par les formules (4). Par conséquent, l'hypothèse dim $M < 2n$ est indispensable dans ces cas pour la validité du théorème 1. Réciproquement nous allons maintenant montrer que l'existence d'une telle transformation χ_0 permet d'obtenir la thèse du théorème 1 sans aucune hypothèse sur la dimension de l'espace M.

THÉORÈME 3. *L'existence d'une fonction $\chi_0 \in S_n^{S_n \times S_n}$, transformant pour un certain $a \in S_n$ chacune des surfaces sphériques $S_n \times (a)$ et $(a) \times S_n$ en S_n avec le grade 1 entraîne que la h-somme existe pour tout couple des transformations continues d'un espace arbitraire M en S_n.*

DÉMONSTRATION. Étant données deux transformations $\varphi, \psi \in S_n^M$, posons:

$$\bar\varphi(x) = (\varphi(x), a); \qquad \bar\psi(x) = (a, \psi(x)); \qquad \theta(x) = (\varphi(x), \psi(x))$$

et prouvons au préalable que pour tout vrai cycle n-dimensionnel $\mathfrak{C}$ de M le vrai cycle $\mathfrak{C}_\theta$ est homologue dans $S_n \times S_n$ au vrai cycle $\mathfrak{C}_{\bar\varphi} + \mathfrak{C}_{\bar\psi}$. Dans ce but envisageons un cube euclidien Q_{n+1} à $n + 1$ dimensions, dont nous pouvons identifier la frontière avec S_n. La frontière du cube $Q_{2n+2} = Q_{n+1} \times Q_{n+1}$ est, d'une part, égale à la somme des ensembles $Q_{n+1} \times S_n$ et $S_n \times Q_{n+1}$, dont la partie commune est $S_n \times S_n$ et, d'autre part, topologiquement identique avec la surface sphérique S_{2n+1}. L'existence d'une déformation continue de Q_{n+1} en lui-même au point a entraîne que la fonction $\theta(x)$ est, dans l'espace $(Q_{n+1} \times S_n)^M$, homotope (et par conséquent aussi homologue) à la fonction $\bar\psi(x)$ et dans l'espace $(S_n \times Q_{n+1})^M$, à la fonction $\bar\varphi(x)$. Il en résulte que le vrai cycle $\mathfrak{C}' = \mathfrak{C}_\theta - (\mathfrak{C}_{\bar\varphi} + \mathfrak{C}_{\bar\psi})$ est dans $Q_{n+1} \times S_n$ homologue au vrai cycle $\mathfrak{C}_{\bar\varphi} - \mathfrak{C}_{\bar\varphi} = 0$. Pareillement $\mathfrak{C}'$ est homologue à 0 dans $S_n \times Q_{n+1}$. En tenant compte du fait que chaque vrai cycle de dimension $n + 1$ est homologue à 0 dans $S_{2n+1} = (Q_{n+1} \times S_n) + (S_n + Q_{n+1})$, on en conclut[14] que $\mathfrak{C}'$ est homologue à 0 dans l'ensemble $S_n \times S_n$, c.à d. $\mathfrak{C}_\theta \sim \mathfrak{C}_{\bar\varphi} + \mathfrak{C}_{\bar\psi}$ dans $S_n \times S_n$.

Ceci établi nous allons montrer que la fonction $\chi \in S_n^M$ définie par la formule

$$\chi(x) = \chi_0[\varphi(x), \psi(x)]$$

[13] Voir H. Hopf, *Über die Abbildungen von Sphären auf Sphären niedriger Dimension*, Fund. Math. 25 (1935), p. 436, th. V.

[14] Nous nous appuyons ici sur le simple theoreme suivant: Si M est une somme de deux sous-ensembles compacts M_1 et M_2, où $M_1 \cdot M_2$ contient un vrai cycle n-dimensionnel homologue à 0 dans chacun des ensembles M_i, mais non homologue à 0 dans $M_1 \cdot M_2$, alors il existe un vrai cycle de dimension $n + 1$ non homologue à 0 dans M. Voir par exemple la note citée dans le renvoi[12], p. 13.

est une h-somme de φ et ψ. Il faut démontrer que pour tout vrai cycle $\mathfrak{C}$ de M on a $\mathfrak{C}_\chi \sim \mathfrak{C}_\varphi + \mathfrak{C}_\psi$ dans S_n. La surface sphérique S_n étant acyclique en toutes les dimensions $\neq n$, nous pouvons ne considérer que de vrais cycles de dimension n. Comme nous l'avons établi, $\mathfrak{C}_\theta$ est dans ce cas homologue dans $S_n \times S_n$ au vrai cycle $\mathfrak{C}_{\bar\varphi} + \mathfrak{C}_{\bar\psi}$ qui est transformé par χ_0 en vrai cycle $(\mathfrak{C}_{\bar\varphi})_{\chi_0} + (\mathfrak{C}_{\bar\psi})_{\chi_0}$. Or, la fonction χ_0 transformant chacune des surfaces sphériques $S_n \times (a)$ et $(a) \times S_n$ en S_n avec le grade 1, le vrai cycle $(\mathfrak{C}_{\bar\varphi})_{\chi_0}$ est homologue dans S_n à $\mathfrak{C}_\varphi$ et le vrai cycle $(\mathfrak{C}_{\bar\psi})_{\chi_0}$ à $\mathfrak{C}_\psi$. Il en résulte $\mathfrak{C}_\chi = (\mathfrak{C}_\theta)_{\chi_0} \sim (\mathfrak{C}_{\bar\varphi})_{\chi_0} + (\mathfrak{C}_{\bar\psi})_{\chi_0} \sim \mathfrak{C}_\varphi + \mathfrak{C}_\psi$, c.q.f.d.

6. En tenant compte du fait qu'il existe dans le cas $n = 1, 3, 7$ une transformation χ_0 remplissant les hypothèses du Numéro précédent,[15] on tire du théorème 3:

Corollaire. *Les types d'homologie des transformations d'un espace M en S_n constituent pour $n = 1, 3, 7$ un groupe abélien dans lequel la h-sommation joue le rôle de l'addition.*

Varsovie.

---

[15] Voir H. Hopf, l.c. p. 436, th. VI.

Annals of Mathematics
Vol. 38, No. 3, July, 1937

QUARTIC FIELDS WITH THE SYMMETRIC GROUP

By D. M. Dribin[1]

(Received December 8, 1936)

1. **Introduction.** As is well known in the theory of algebraic number fields, a prime ideal $\mathfrak{p}$ in a field k characterizes certain subgroups of the Galois group of a normal field $K \mid k$—the so-called groups of decomposition, of inertia, and of ramification. These subgroups describe the behavior of the prime ideal $\mathfrak{p}$ in K and in certain subfields of K. For a given $K \mid k$, well-known results in algebraic number theory describe this behavior, and, in particular, if $K \mid k$ is abelian, the class-field theory throws further light on the indices of ramification, on the conductor of $K \mid k$ and the like.

This paper is concerned with a converse problem. We shall consider the possible types of groups of decomposition, inertia, and ramification connected with the normal extension of a quartic field over R, the field of the rational numbers, having the symmetric group, and determine which of these types may exist.

This problem is one that has been already treated by Hasse[2] for the case of cubic fields with the symmetric group, and, as a matter of fact, our work is intimately connected with his, as we shall soon see. Other work of this same nature in which class-field theoretic methods were employed has been done by Porusch,[3] who considered special metabelian fields, and by Rosenblüth,[4] who studied tha "quaternionic" field.

2. **Notations. Construction of the tables.** For what follows we shall need a complete list of the subgroups of the symmetric group of degree 4, and shall, in addition, represent certain of these subgroups by permutations on 4 letters so that the interrelations of the various groups may thereby be made more clear.

The complete list (excluding the unit element) is:

---

[1] National Research Fellow.

[2] Hasse, *Arithmetische Theorie der kubischen Zahlkörper auf klassenkörpertheoretischer Grundlage*, *Math. Zeitschr.*, 31 (1930), pp. 565–582.

[3] *Die Arithmetik in Zahlkörpern, deren zugehörige Galoissche Körper spezielle metabelsche Gruppen besitzen, auf klassenkörpertheoretischer Grundlage*, *Math. Zeitschr.*, 37 (1933), pp. 134–160.

[4] *Die arithmetische Theorie und die Konstruktion der Quaternionenkörper auf klassenkörpertheoretischer Grundlage*, *Monatsh. f. Math. u. Physik*, 41 (1934), pp. 85–125.

S, the symmetric group itself,

A, the alternating group of degree 4,

G_8: $\{1, (12), (34), (12)(34), (13)(24), (14)(23), (1324), (1423)\}$,

G_8', G_8'', the conjugates of G_8 in S,

G_6 : $\{1, (12), (13), (23), (123), (132)\}$,

G_6', G_6'', G_6''', the conjugates of G_6 in S,

(1) V_4 : $\{1, (12)(34), (13)(24), (14)(23)\}$ (Vierergroup),

C_4 : $\{1, (1324), (1423), (12)(34)\}$,

C_4', C_4'', the conjugates of C_4 in S,

G_4 : $\{1, (12), (34), (12)(34)\}$,

G_4', G_4'', the conjugates of G_4 in S,

C_3 : $\{1, (123), (132)\}$; C_3', C_3'', C_3''', the conjugates of C_3 in S,

V_2 : $\{1, (12)(34)\}$; V_2', V_2'' the conjugates of V_2 in S,

C_2 : $\{1, (12)\}$; C_2', $\cdots$, $C_2^{(v)}$ the conjugates of C_2 in S.

If, now, N is a normal field with the symmetric group of degree 4, there corresponds to the preceding list of subgroups of S, a list of the various subfields of N over R. We shall mention only

N, normal of degree 24 $\leftrightarrow 1$,

Λ_1, Λ_2, Λ_3, conjugate fields of degree 12 $\leftrightarrow V_2, V_2', V_2''$,

(2) B, normal sextic field $\leftrightarrow V_4$,

K, K_1, K_2, K_3, conjugate quartic fields $\leftrightarrow G_6, G_6', G_6'', G_6'''$,

$\Omega, \Omega_1, \Omega_2$, conjugate cubic fields $\leftrightarrow G_8, G_8', G_8''$,

Q, quadratic field $\leftrightarrow A$.

The groups associated with each of the fields, as is indicated above, are those subgroups associated by the fundamental theorem of the Galois theory. In particular, we see also that the group of $N \mid B$ is V_4, that the group of $B \mid R$ is isomorphic to the symmetric group on three letters, and that B is a normal extension of the non-normal cubic field Ω. It is now clear how the work of Hasse, previously referred to, comes into the picture.

We now explain the construction of the various tables (pp. 744–746). We denote by G_Z, G_T, and the G_{V_i}, for a given rational prime p, the groups of decomposition, inertia, and ramification, respectively, in N. Table A yields information concerning those primes p which are not divisors of the discriminant of $N \mid R$, i.e., for which $G_T = 1$; Table B yields information concerning the "regular" prime divisors of the discriminant, i.e., for which $G_T \neq 1$, $G_{V_1} = 1$; and Table C yields information concerning the "irregular" prime divisors of

the discriminant, i.e., for which $G_{V_1} \neq 1$. Table D contains all[5] the possible cases that, though belonging properly to Table C, will be proved to be impossible of realization.

The possible types of the Hilbert subgroup series are briefly described as follows.[6] We have always

$$S \supseteq G_Z \supseteq G_T \supseteq G_{V_{v_1}} \supseteq \cdots \supseteq G_{V_t} = 1,$$

where $G_{V_1}, \cdots, G_{V_{v_1}}$, are all equal, $G_{V_{v_1+1}}, \cdots, G_{V_{v_2}}$ are all equal and proper subgroups of G_{V_1}, and so on. The groups in the Hilbert series from G_T on are normal divisors of G_Z, the factor groups G_Z/G_T and G_T/G_{V_1} are cyclic, and the factor groups $G_{V_i}/G_{V_{i+1}}$ are all abelian of type $(p, p, \cdots, p)$. The orders of the G_{V_i} are powers of p and G_T/G_{V_1} has order e_0 prime to p. These considerations yield the types of Hilbert series that are listed in column 1 of Tables A-D.[7]

Assume that the orders of S/G_Z, G_Z/G_T, and G_T are g, f, and $e = e_0 p^{R_1}$, respectively. Then p decomposes in N in the following manner:

$$p = (\mathfrak{P}_1 \cdots \mathfrak{P}_g)^e, \qquad N_{N \mid R}(\mathfrak{P}_i) = p^f.$$

Also $p^f \equiv 1 \pmod{e_0}$. These facts yield columns 2 and 4 of the tables. That column 2 often gives finer results than are obtained by these means is due to the fact that we are combining the results of our tables with those in the table in Hasse's paper. Thus (6, 5)[8] of Hasse's table yields only $p \neq 2$ whereas in $(B, 3, 2)$*—a corresponding case—we get $p \equiv 1 \pmod 4$.

The groups in the Hilbert series for any subfield Σ of N are obtained by intersecting the groups in the Hilbert series for N with that group in the Galois series for N which corresponds to Σ. Thus, if $\Sigma = B$, and if V_4, $\bar{G}_Z$, $\bar{G}_T$, $\bar{G}_{V_i}$ are the corresponding groups, and if the orders of $V_4/\bar{G}_Z$, $\bar{G}_Z/\bar{G}_T$, and $\bar{G}_T$ are $\bar{g}$, $\bar{f}$, and $\bar{e}$, respectively, and if $\mathfrak{p}$ is a prime ideal divisor of p in B, then

$$\mathfrak{p} = (\mathfrak{P}_1 \cdots \mathfrak{P}_{\bar{g}})^{\bar{e}}, \qquad N_{N \mid B}(\mathfrak{P}_i) = \mathfrak{p}^{\bar{f}}.$$

To find the decomposition of p in B, we use the well known formulas

$$\bar{e}e' = e, \qquad \bar{f}f' = f, \qquad \bar{g}g' = g,$$

---

[5] That is, all except two that will be disposed of in a footnote a little further on.

[6] For proofs of the various statements that are being asserted here and in the succeeding paragraphs, the reader is referred to Hasse's *Bericht, Ia, Jahresber. d. Deutsch. Math.-Ver.* 36 (1927), §§8, 9.

[7] It may be said that further cases would arise were we to choose G_8', say, instead of G_8, but such cases would yield nothing new; their effect would be to permute the subfields of N. That is, instead of using K or Ω, fields *conjugate* to K or Ω would be used.

[8] (6, 5) will be used to indicate the entry in the sixth *row* and fifth *column* of Hasse's table; $(B, 3, 2)$ will be used to indicate the entry in the third *row* and second *column* of our table B whereas $(B, 3)$ will simply refer to the case corresponding to the third *row* in table B. We shall continue to employ this convenient notation.

S, the symmetric group itself,

A, the alternating group of degree 4,

G_8: $\{1, (12), (34), (12)(34), (13)(24), (14)(23), (1324), (1423)\}$,

G_8', G_8'', the conjugates of G_8 in S,

G_6 : $\{1, (12), (13), (23), (123), (132)\}$,

G_6', G_6'', G_6''', the conjugates of G_6 in S,

(1) V_4 : $\{1, (12)(34), (13)(24), (14)(23)\}$ (Vierergroup),

C_4 : $\{1, (1324), (1423), (12)(34)\}$,

C_4', C_4'', the conjugates of C_4 in S,

G_4 : $\{1, (12), (34), (12)(34)\}$,

G_4', G_4'', the conjugates of G_4 in S,

C_3 : $\{1, (123), (132)\}$; C_3', C_3'', C_3''', the conjugates of C_3 in S,

V_2 : $\{1, (12)(34)\}$; V_2', V_2'' the conjugates of V_2 in S,

C_2 : $\{1, (12)\}$; C_2', $\cdots$, $C_2^{(v)}$ the conjugates of C_2 in S.

If, now, N is a normal field with the symmetric group of degree 4, there corresponds to the preceding list of subgroups of S, a list of the various subfields of N over R. We shall mention only

	N, normal of degree 24	$\leftrightarrow 1$,
	Λ_1, Λ_2, Λ_3, conjugate fields of degree 12	$\leftrightarrow V_2$, V_2', V_2'',
(2)	B, normal sextic field	$\leftrightarrow V_4$,
	K, K_1, K_2, K_3, conjugate quartic fields	$\leftrightarrow G_6$, G_6', G_6'', G_6''',
	$\Omega, \Omega_1, \Omega_2$, conjugate cubic fields	$\leftrightarrow G_8$, G_8', G_8'',
	Q, quadratic field	$\leftrightarrow A$.

The groups associated with each of the fields, as is indicated above, are those subgroups associated by the fundamental theorem of the Galois theory. In particular, we see also that the group of $N \mid B$ is V_4, that the group of $B \mid R$ is isomorphic to the symmetric group on three letters, and that B is a normal extension of the non-normal cubic field Ω. It is now clear how the work of Hasse, previously referred to, comes into the picture.

We now explain the construction of the various tables (pp. 744–746). We denote by G_Z, G_T, and the G_{V_i}, for a given rational prime p, the groups of decomposition, inertia, and ramification, respectively, in N. Table A yields information concerning those primes p which are not divisors of the discriminant of $N \mid R$, i.e., for which $G_T = 1$; Table B yields information concerning the "regular" prime divisors of the discriminant, i.e., for which $G_T \neq 1$, $G_{V_1} = 1$; and Table C yields information concerning the "irregular" prime divisors of

the discriminant, i.e., for which $G_{V_1} \neq 1$. Table D contains all[5] the possible cases that, though belonging properly to Table C, will be proved to be impossible of realization.

The possible types of the Hilbert subgroup series are briefly described as follows.[6] We have always

$$S \supseteq G_Z \supseteq G_T \supseteq G_{V_{v_1}} \supseteq \cdots \supseteq G_{V_t} = 1,$$

where $G_{V_1}, \cdots, G_{V_{v_1}}$, are all equal, $G_{V_{v_1+1}}, \cdots, G_{V_{v_2}}$ are all equal and proper subgroups of G_{V_1}, and so on. The groups in the Hilbert series from G_T on are normal divisors of G_Z, the factor groups G_Z/G_T and G_T/G_{V_1} are cyclic, and the factor groups $G_{V_i}/G_{V_{i+1}}$ are all abelian of type $(p, p, \cdots, p)$. The orders of the G_{V_i} are powers of p and G_T/G_{V_1} has order e_0 prime to p. These considerations yield the types of Hilbert series that are listed in column 1 of Tables A-D.[7]

Assume that the orders of S/G_Z, G_Z/G_T, and G_T are g, f, and $e = e_0 p^{R_1}$, respectively. Then p decomposes in N in the following manner:

$$p = (\mathfrak{P}_1 \cdots \mathfrak{P}_g)^e, \qquad N_{N \mid R}(\mathfrak{P}_i) = p^f.$$

Also $p^f \equiv 1 \pmod{e_0}$. These facts yield columns 2 and 4 of the tables. That column 2 often gives finer results than are obtained by these means is due to the fact that we are combining the results of our tables with those in the table in Hasse's paper. Thus (6, 5)[8] of Hasse's table yields only $p \neq 2$ whereas in $(B, 3, 2)^*$—a corresponding case—we get $p \equiv 1 \pmod 4$.

The groups in the Hilbert series for any subfield Σ of N are obtained by intersecting the groups in the Hilbert series for N with that group in the Galois series for N which corresponds to Σ. Thus, if $\Sigma = B$, and if V_4, $\bar{G}_Z$, $\bar{G}_T$, $\bar{G}_{V_i}$ are the corresponding groups, and if the orders of $V_4/\bar{G}_Z$, $\bar{G}_Z/\bar{G}_T$, and $\bar{G}_T$ are $\bar{g}$, $\bar{f}$, and $\bar{e}$, respectively, and if $\mathfrak{p}$ is a prime ideal divisor of p in B, then

$$\mathfrak{p} = (\mathfrak{P}_1 \cdots \mathfrak{P}_{\bar{g}})^{\bar{e}}, \qquad N_{N \mid B}(\mathfrak{P}_i) = \mathfrak{p}^{\bar{f}}.$$

To find the decomposition of p in B, we use the well known formulas

$$\bar{e}e' = e, \qquad \bar{f}f' = f, \qquad \bar{g}g' = g,$$

[5] That is, all except two that will be disposed of in a footnote a little further on.

[6] For proofs of the various statements that are being asserted here and in the succeeding paragraphs, the reader is referred to Hasse's *Bericht, Ia, Jahresber. d. Deutsch. Math.-Ver.* 36 (1927), §§8, 9.

[7] It may be said that further cases would arise were we to choose G_8', say, instead of G_8, but such cases would yield nothing new; their effect would be to permute the subfields of N. That is, instead of using K or Ω, fields *conjugate* to K or Ω would be used.

[8] (6, 5) will be used to indicate the entry in the sixth *row* and fifth *column* of Hasse's table; $(B, 3, 2)$ will be used to indicate the entry in the third *row* and second *column* of our table B whereas $(B, 3)$ will simply refer to the case corresponding to the third *row* in table B. We shall continue to employ this convenient notation.

where g' is the number of prime ideal divisors of p in B, f' their common degree, and e' their common order. Thus we get[9] column 5 of the tables and, in a similar manner, column 9. If we represent by d the discriminant of $Q \mid R$ then column 14 merely yields the well-known criterion describing the decomposition of a prime in a quadratic field.

Column 6 will follow if we consider B as the ground field and investigate the decomposition of a prime ideal $\mathfrak{p}$ in B in the quadratic fields Λ_1 and Λ_2 over B.

To obtain column 7, we proceed as above but use the rules[10]

$$\sum_{i=0}^{G-1} E_i F_i = 4, \quad E_i \bar{E}_i = e, \quad F_i \bar{F}_i = f, \quad \sum_{i=0}^{G-1} G_i = g, \quad \bar{E}_i \bar{F}_i \bar{G}_i = 6,$$

where G is the number of prime ideal divisors of p in K, E_i their respective degrees and E_i their respective orders. G_i, F_i, E_i are the corresponding numbers of $N \mid K$. In the same way (except that we use the numbers 3 and 8 in place of 4 and 6 above) we get column 8.

If D_N represents the discriminant of $N \mid R$ and if $(D_N)_p$ represents the highest power of p contained in D_N, then

$$(D_N)_p = p^{fg[(e-1)+(p^{R_1}-1)v_1+(p^{R_2}-1)(v_2-v_1)+\cdots]},$$

where p^{R_1} is the order of $G_{V_1}, \cdots, G_{V_{v_1}}$, p^{R_2} that of $G_{V_{v_1+1}}, \cdots, G_{V_{v_2}}$, etc. Thus we get column 10.

Let $\mathfrak{p}_1, \cdots, \mathfrak{p}_g$ be the prime ideal divisors of p in B. Then

$$(D_{N \mid B})_p = (\mathfrak{p}_1 \cdots \mathfrak{p}_g)^{gf[(e-1)+(p^{R_1}-1)v_1+(p^{R_2}-1)(v_2-v_1)+\cdots]},$$

where the notation refers to the Hilbert subgroup series for B. We then have

$$(D_N)_p = (D_B)_p^4 \, \dot{N}_{B \mid R} (D_{N \mid B})_p .$$

The second factor on the right can be easily determined by means of column 5.

To find $(D_Q)_p$, $(D_\Omega)_p$, $(D_K)_p$, we proceed in a similar manner, using the facts

$$(D_N)_p = (D_Q)_p^{12} N(D_{N \mid Q})_p = (D_\Omega)_p^8 N(D_{N \mid \Omega})_p = (D_K)_p^6 N(D_{N \mid K})_p .$$

Column 15 yields the contribution of p to the conductors of the ideal groups

---

[9] At this point we can easily dispose of two cases which were not listed in Table D for purposes of unity. In the first case, $G = S$, $G_Z = G_T = A$, $G_{V_1} = V_4$, $G_{V_1+1} = 1$; hence in B, $p = (\mathfrak{p}_1 \mathfrak{p}_2)^3$. But according to Hasse's table, p must be 3 or $\equiv 1 \pmod 3$, whereas we must have $p = 2$.

In the second case, $G = G_Z = S$, $G_T = G_{V_1} = V_4$, $G_{V_1+1} = 1$; hence in B, $p = \mathfrak{p}$. But by Hasse's table such a case cannot exist.

[10] $i = 0$ is to be interpreted as the absence of a subscript. In such fields as K and Ω, there will be a "preference" of a certain prime ideal divisior of p in K or Ω but this preference depends on our choice of K from its set of four conjugates, or of Ω from its set of conjugate fields. See footnote on page 741.

in B to which the Λ_i are the class fields. By the class-field theory (since the Λ_i are quadratic over B), these conductors are equal to the discriminants of the $\Lambda_i \mid B$ and the latter quantities are easily obtained.

It remains only to discuss the determination of the indices of ramification that are listed in column 3. These values (corresponding to existent cases) are gleaned from Table E, which gives all of the possible values for (v_1, v_2) according to the considerations that we shall now make.

We employ the following known facts. If $K \mid k$ is cyclic of prime degree l and if v is the number of the groups of ramification in K of a prime ideal divisor $\mathfrak{l}$ of l, then

$$1 \leqq v \leqq \frac{el}{l-1}, \qquad \text{with } v = \frac{el}{l-1} \qquad \text{or } (v, l) = 1,$$

where e is the order of $\mathfrak{l}$ in l. Furthermore, if k contains all the l^{th} roots of unity, then $e^* = \dfrac{e}{l-1}$ is an integer and the exact power of $\mathfrak{l}$ that divides $\mathfrak{F}_{K \mid k}$, the conductor of $K \mid k$, is[11] $\mathfrak{l}^{e+e^*+1}$.

Finally, we need a result of Speiser[12] which states that if p is the prime dividing the orders of a series of groups of ramification in a normal field $K \mid k$, and that if v_i and v_j are two indices of ramification, then $v_i \equiv v_j \pmod{p}$.

We employ these theorems by noting that $N \mid \Lambda_i$ is quadratic and that $N \mid \bar{\Lambda}$ is cyclic of degree 3 (where $\bar{\Lambda}$ corresponds to C_3 in the Galois series for N). The results are given in Table E.

3. **The irregular prime divisor** 2. In this section we shall study the cases in Table C arising from the irregular prime divisor of the discriminant of $N \mid R$, two. We must first learn to what power a prime ideal divisor $\mathfrak{p}$ of 2 may occur in the discriminant of a relatively quadratic field $\Lambda \mid B$, where $\Lambda = B(\sqrt{\mu})$ and μ is an integer in B.

We may assume that if $\mathfrak{p}$ is a prime ideal divisor of 2 that μ is divisible by at most $\mathfrak{p}^1$. We write $\mu = (\mathfrak{p}^{2m})\mathfrak{q}$, where $\mathfrak{q}$ is divisible by at most $\mathfrak{p}^1$, $m = 2^{m'}\bar{m}$, $h = 2^{h'}\bar{h}$, where h is the (absolute) class number of B, and $\bar{m}\bar{h}$ is odd. If $m' \geqq h'$, then

$$\mu^{\bar{h}} = (\mathfrak{p}^{2^{m'-h'}\bar{m}})^{2h}\mathfrak{q}^{\bar{h}} = \omega_1^2\omega_2, \qquad (\omega_i \text{ in } B),$$

and $\sqrt{\omega_2} = \omega_1 \mu^{\frac{1}{2}(\bar{h}-1)}\sqrt{\mu}$, so that $B(\sqrt{\omega_2}) = B(\sqrt{\mu})$ and ω_2 is divisible by at most $\mathfrak{p}^1$.

If $m' < h'$, we can determine an integer π in B such that[13] $\mathfrak{p}^{2^{m'}\bar{m}(2^{h'-m'}-1)} \parallel \pi$. Then

$$\bar{\mu} = \mu\pi^2 = (\mathfrak{p}^{2^{h'}\bar{m}})^2\mathfrak{r},$$

---

[11] Hasse, *Bericht, Teil II*, §9.

[12] Speiser, *Die Zerlegungsgruppe*, *J. f. Math. u. Physik*, 149 (1919), p. 182.

[13] By $\mathfrak{p}^\alpha \parallel \mathfrak{m}$ is meant: $\mathfrak{p}^\alpha \mid \mathfrak{m}$, and $\mathfrak{p}^{\alpha+1} \mid \mathfrak{m}$ is false.

TABLE

	1					2	3	4		5		6		7
	G	G_Z	G_T	G_{V_v}	$G_{V_{v+1}}$	Conditions on p	v	In N, $p =$	Degree	In B, $p =$	Degree	In Λ_1, $\mathfrak{p} =$	In Λ_2, $\mathfrak{p} =$	K, $p =$
1	S	1	1	1	1	–	–	$\mathfrak{P}_1 \cdots \mathfrak{P}_{24}$	1	$\mathfrak{p}_1 \cdots \mathfrak{p}_6$	1	$\mathfrak{p}_1'\mathfrak{p}_2'$	$\mathfrak{p}_1''\mathfrak{p}_2''$	$\cdots \bar{\mathfrak{P}}_4$
2	S	C_4	1	1	1	–	–	$\mathfrak{P}_1 \cdots \mathfrak{P}_6$	4	$\mathfrak{p}_1\mathfrak{p}_2\mathfrak{p}_3$	2	$\mathfrak{p}_1'\mathfrak{p}_2'$	$\mathfrak{p}''$	$\bar{\mathfrak{P}}$
3	S	C_3	1	1	1	–	–	$\mathfrak{P}_1 \cdots \mathfrak{P}_8$	3	$\mathfrak{p}_1\mathfrak{p}_2$	3	$\mathfrak{p}_1'\mathfrak{p}_2'$	$\mathfrak{p}_1''\mathfrak{p}_2''$	$\bar{\mathfrak{P}}_1\bar{\mathfrak{P}}_2$
4	S	V_2	1	1	1	–	–	$\mathfrak{P}_1 \cdots \mathfrak{P}_{12}$	2	$\mathfrak{p}_1 \cdots \mathfrak{p}_6$	1	$\mathfrak{p}_1'\mathfrak{p}_2'$	$\mathfrak{p}''$	$\bar{\mathfrak{P}}_1\bar{\mathfrak{P}}_2$
5	S	C_2	1	1	1	–	–	$\mathfrak{P}_1 \cdots \mathfrak{P}_{12}$	2	$\mathfrak{p}_1\mathfrak{p}_2\mathfrak{p}_3$	2	$\mathfrak{p}_1'\mathfrak{p}_2'$	$\mathfrak{p}_1''\mathfrak{p}_2''$	$\bar{\mathfrak{P}}_1\bar{\mathfrak{P}}_2\bar{\mathfrak{P}}_3$

TABLE B

	G	G_Z	G_T	G_{V_v}	$G_{V_{v+1}}$	Conditions on p	v	In N, $p =$	Degree	In B, $p =$	Degree	In Λ_1, $\mathfrak{p} =$	In Λ_2, $\mathfrak{p} =$	K, $p =$
1	S	G_6	C_3	1	1	$p \equiv -1(3)$	–	$(\mathfrak{P}_1 \cdots \mathfrak{P}_4)^3$	2	$\mathfrak{p}^3$	2	$\mathfrak{p}_1'\mathfrak{p}_2'$	$\mathfrak{p}_1''\mathfrak{p}_2''$	$\bar{\mathfrak{P}}_1\bar{\mathfrak{P}}_2^3$
2	S	V_4	V_2	1	1	$p \neq 2$	–	$(\mathfrak{P}_1 \cdots \mathfrak{P}_6)^2$	2	$\mathfrak{p}_1 \cdots \mathfrak{p}_6$	1	$\mathfrak{p}'$	$\mathfrak{p}''^2$	$\bar{\mathfrak{P}}^2$
3	S	C_4	C_4	1	1	$p \equiv 1$ (4)	–	$(\mathfrak{P}_1 \cdots \mathfrak{P}_6)^4$	1	$(\mathfrak{p}_1\mathfrak{p}_2\mathfrak{p}_3)^2$	1	$\mathfrak{p}_1'\mathfrak{p}_2'$	$\mathfrak{p}''^2$	$\bar{\mathfrak{P}}^4$
4	S	V_2	V_2	1	1	$p \neq 2$	–	$(\mathfrak{P}_1 \cdots \mathfrak{P}_{12})^2$	1	$\mathfrak{p}_1 \cdots \mathfrak{p}_6$	1	$\mathfrak{p}_1'\mathfrak{p}_2'$	$\mathfrak{p}''^2$	$(\bar{\mathfrak{P}}_1\bar{\mathfrak{P}}_2)^2$
5	S	C_2	C_2	1	1	$p \neq 2$	–	$(\mathfrak{P}_1 \cdots \mathfrak{P}_{12})^2$	1	$(\mathfrak{p}_1\mathfrak{p}_2\mathfrak{p}_3)^2$	1	$\mathfrak{p}_1'\mathfrak{p}_2'$	$\mathfrak{p}_1''\mathfrak{p}_2''$	$\bar{\mathfrak{P}}_1\bar{\mathfrak{P}}_2\bar{\mathfrak{P}}_3^2$
6	S	C_3	C_3	1	1	$p \equiv +1(3)$	–	$(\mathfrak{P}_1 \cdots \mathfrak{P}_8)^3$	1	$(\mathfrak{p}_1\mathfrak{p}_2)^3$	1	$\mathfrak{p}_1'\mathfrak{p}_2'$	$\mathfrak{p}_1''\mathfrak{p}_2''$	$\bar{\mathfrak{P}}_1\bar{\mathfrak{P}}_2^3$
7	S	C_4	V_2	1	1	$p \neq 2$	–	$(\mathfrak{P}_1 \cdots \mathfrak{P}_6)^2$	2	$\mathfrak{p}_1\mathfrak{p}_2\mathfrak{p}_3$	2	$\mathfrak{p}_1'\mathfrak{p}_2'$	$\mathfrak{p}''^2$	$\bar{\mathfrak{P}}^2$
8	S	G_8	V_2	1	1	$p \neq 2$	–	$(\mathfrak{P}_1\mathfrak{P}_2\mathfrak{P}_3)^2$	4	$\mathfrak{p}_1\mathfrak{p}_2\mathfrak{p}_3$	2	$\mathfrak{p}'$	$\mathfrak{p}''^2$	$\bar{\mathfrak{P}}^2$
9	S	G_4	V_2	1	1	$p \neq 2$	–	$(\mathfrak{P}_1 \cdots \mathfrak{P}_6)^2$	2	$\mathfrak{p}_1\mathfrak{p}_2\mathfrak{p}_3$	2	$\mathfrak{p}_1'\mathfrak{p}_2'$	$\mathfrak{p}''^2$	$(\bar{\mathfrak{P}}_1\bar{\mathfrak{P}}_2)^2$
10	S	G_4	C_2	1	1	$p \neq 2$	–	$(\mathfrak{P}_1 \cdots \mathfrak{P}_6)^2$	2	$(\mathfrak{p}_1\mathfrak{p}_2\mathfrak{p}_3)^2$	1	$\mathfrak{p}_1'\mathfrak{p}_2'$	$\mathfrak{p}''$	$\bar{\mathfrak{P}}_1\bar{\mathfrak{P}}_2^2$

TABLE C

	G	G_Z	G_T	G_{V_v}	$G_{V_{v+1}}$	Conditions on p	v	In N, $p =$	Degree	In B, $p =$	Degree	In Λ_1, $\mathfrak{p} =$	In Λ_2, $\mathfrak{p} =$	K, $p =$
1	S	S	A	V_4	1	$p = 2$	1, 3	$\mathfrak{P}^{12}$	2	$\mathfrak{p}^3$	2	$\mathfrak{p}'^2$	$\mathfrak{p}''^2$	$\bar{\mathfrak{P}}^4$
2	S	G_6	G_6	C_3	1	$p = 3$	1, 3; $v = 1$ only when $d \equiv 3(9)$	$(\mathfrak{P}_1 \cdots \mathfrak{P}_4)^6$	1	$\mathfrak{p}^6$	1	$\mathfrak{p}_1'\mathfrak{p}_2'$	$\mathfrak{p}_1''\mathfrak{p}_2''$	$\bar{\mathfrak{P}}_1\bar{\mathfrak{P}}_2^3$
3	S	G_6	C_3	C_3	1	$p = 3$	1	$(\mathfrak{P}_1 \cdots \mathfrak{P}_4)^3$	2	$\mathfrak{p}^3$	2	$\mathfrak{p}_1'\mathfrak{p}_2'$	$\mathfrak{p}_1''\mathfrak{p}_2''$	$\bar{\mathfrak{P}}_1\bar{\mathfrak{P}}_2^3$
4	S	V_4	V_2	V_2	1	$p = 2$	2	$(\mathfrak{P}_1 \cdots \mathfrak{P}_6)^2$	2	$\mathfrak{p}_1 \cdots \mathfrak{p}_6$	1	$\mathfrak{p}'$	$\mathfrak{p}''^2$	$\bar{\mathfrak{P}}^2$
5	S	C_3	C_3	C_3	1	$p = 3$	1	$(\mathfrak{P}_1 \cdots \mathfrak{P}_8)^3$	1	$(\mathfrak{p}_1\mathfrak{p}_2)^3$	1	$\mathfrak{p}_1'\mathfrak{p}_2'$	$\mathfrak{p}_1''\mathfrak{p}_2''$	$\bar{\mathfrak{P}}_1\bar{\mathfrak{P}}_2^3$
6	S	V_2	V_2	V_2	1	$p = 2$	2	$(\mathfrak{P}_1 \cdots \mathfrak{P}_{12})^2$	1	$\mathfrak{p}_1 \cdots \mathfrak{p}_6$	1	$\mathfrak{p}_1'\mathfrak{p}_2'$	$\mathfrak{p}''^2$	$(\bar{\mathfrak{P}}_1\bar{\mathfrak{P}}_2)^2$
7	S	C_2	C_2	C_2	1	$p = 2$	1, 2	$(\mathfrak{P}_1 \cdots \mathfrak{P}_{12})^2$	1	$(\mathfrak{p}_1\mathfrak{p}_2\mathfrak{p}_3)^2$	1	$\mathfrak{p}_1'\mathfrak{p}_2'$	$\mathfrak{p}_1''\mathfrak{p}_2''$	$\bar{\mathfrak{P}}_1\bar{\mathfrak{P}}_2\bar{\mathfrak{P}}_3^2$

TABLE A

,	6	7		8		9		10	11	12	13	14	15		
	In Λ_2 $\mathfrak{p}$ =	K, p =	Degrees	In Ω, p =	Degrees	In Q, p =	Degree	$(D_N)_p$	$(D_B)_p$	$(D_\Omega)_p$	d_p	$\left(\frac{d}{p}\right)$	$\mathfrak{F}_p^{(1)}$	$\mathfrak{F}_p^{(2)}$	
	$\mathfrak{p}_1''\mathfrak{p}_2'$	$\ldots\bar{\mathfrak{P}}_4$	1, 1, 1, 1	$\mathfrak{p}\,\mathfrak{p}_2\mathfrak{p}_3$	1, 1, 1	$\bar{\mathfrak{p}}_1\bar{\mathfrak{p}}_2$	1	1	1	1	1	+1	1	1	1
	$\mathfrak{p}''$	$\bar{\mathfrak{P}}$	4	$\mathfrak{p}_1\mathfrak{p}_2$	1, 2	$\bar{\mathfrak{p}}$	2	1	1	1	1	−1	1	1	2
	$\mathfrak{p}_1''\mathfrak{p}_2'$	$\bar{\mathfrak{P}}_1\bar{\mathfrak{P}}_2$	1, 3	$\mathfrak{p}$	3	$\bar{\mathfrak{p}}_1\bar{\mathfrak{p}}_2$	1	1	1	1	1	+1	1	1	3
	$\mathfrak{p}''$	$\bar{\mathfrak{P}}_1\bar{\mathfrak{P}}_2$	2, 2	$\mathfrak{p}\,\mathfrak{p}_2\mathfrak{p}_3$	1, 1, 1	$\bar{\mathfrak{p}}_1\bar{\mathfrak{p}}_2$	1	1	1	1	1	+1	1	1	4
	$\mathfrak{p}_1''\mathfrak{p}_2'$	$\bar{\mathfrak{P}}_1\bar{\mathfrak{P}}_2\bar{\mathfrak{P}}_3$	1, 1, 2	$\mathfrak{p}_1\mathfrak{p}_2$	1, 2	$\bar{\mathfrak{p}}$	2	1	1	1	1	−1	1	1	5

TABLE B

	$\mathfrak{p}_1''\mathfrak{p}_2'$	$\bar{\mathfrak{P}}_1\bar{\mathfrak{P}}_2^3$	1, 1	$_3$	1	$\bar{\mathfrak{p}}$	2	p^{16}	p^4	p^2	1	−1	1	1	1
	$\mathfrak{p}''^2$	$\bar{\mathfrak{P}}^2$	2	$_1\;_2\;_3$	1, 1, 1	$\bar{\mathfrak{p}}_1\bar{\mathfrak{p}}_2$	1	p^{12}	1	1	1	+1	1	$\mathfrak{p}$	2
	$\mathfrak{p}''^2$	$\bar{\mathfrak{P}}^4$	1	$_1\;{}^2_2$	1, 1	$\bar{\mathfrak{p}}^2$	1	p^{18}	p^3	p	p	0	1	$\mathfrak{p}$	3
	$\mathfrak{p}''^2$	$(\bar{\mathfrak{P}}_1\bar{\mathfrak{P}}_2)^2$	1, 1	$_1\;_2\;_3$	1, 1, 1	$\bar{\mathfrak{p}}_1\bar{\mathfrak{p}}_2$	1	p^{12}	1	1	1	+1	1	$\mathfrak{p}$	4
	$\mathfrak{p}_1''\mathfrak{p}_2'$	$\bar{\mathfrak{P}}_1\bar{\mathfrak{P}}_2\bar{\mathfrak{P}}_3^2$	1, 1, 1	$_1\;{}^2_2$	1, 1	$\bar{\mathfrak{p}}^2$	1	p^{12}	p^3	p	p	0	1	1	5
	$\mathfrak{p}_1''\mathfrak{p}_2''$	$\bar{\mathfrak{P}}_1\bar{\mathfrak{P}}_2^3$	1, 1	$_3$	1	$\bar{\mathfrak{p}}_1\bar{\mathfrak{p}}_2$	1	p^{16}	p^4	p^2	1	+1	1	1	6
	$\mathfrak{p}''^2$	$\bar{\mathfrak{P}}^2$	2	$_1\;_2$	1, 2	$\bar{\mathfrak{p}}$	2	p^{12}	1	1	1	−1	1	$\mathfrak{p}$	7
	$\mathfrak{p}''^2$	$\bar{\mathfrak{P}}^2$	2	$_1\;_2$	1, 2	$\bar{\mathfrak{p}}$	2	p^{12}	1	1	1	−1	1	$\mathfrak{p}$	8
	$\mathfrak{p}''^2$	$(\bar{\mathfrak{P}}_1\bar{\mathfrak{P}}_2)^2$	1, 1	$_1\;_2$	1, 2	$\bar{\mathfrak{p}}$	2	p^{12}	1	1	1	−1	1	$\mathfrak{p}$	9
	$\mathfrak{p}''$	$\bar{\mathfrak{P}}_1\bar{\mathfrak{P}}_2^2$	2, 1	$_1\;{}^2_2$	1, 1	$\bar{\mathfrak{p}}^2$	1	p^{12}	p^3	p	p	0	1	1	10

TABLE C

	$\mathfrak{p}''^2$	$\bar{\mathfrak{P}}^4$	1	$_3$	1	$\bar{\mathfrak{p}}$	2	$2^{2(11+3v)}$	2^4	2^2	1	−1	$\mathfrak{p}^{1+v}$	$\mathfrak{p}^{1+v}$	1
	$\mathfrak{p}_1''\mathfrak{p}_2''$	$\bar{\mathfrak{P}}_1\bar{\mathfrak{P}}_2^3$	1, 1	$_3$	1	$\bar{\mathfrak{p}}^2$	1	$3^{4(5+2v)}$	3^{5+2v}	3^{2+v}	3	0	1	1	2
	$\mathfrak{p}_1''\mathfrak{p}_2''$	$\bar{\mathfrak{P}}_1\bar{\mathfrak{P}}_2^3$	1, 1	$_3$	1	$\bar{\mathfrak{p}}$	2	3^{32}	3^8	3^4	1	−1	1	1	3
	$\mathfrak{p}''^2$	$\bar{\mathfrak{P}}^2$	2	$_1\;_2\;_3$	1, 1, 1	$\bar{\mathfrak{p}}_1\bar{\mathfrak{p}}_2$	1	2^{36}	1	1	1	+1	1	$\mathfrak{p}^3$	4
	$\mathfrak{p}_1''\mathfrak{p}_2''$	$\bar{\mathfrak{P}}_1\bar{\mathfrak{P}}_2^3$	1, 1	$_3$	1	$\bar{\mathfrak{p}}_1\bar{\mathfrak{p}}_2$	1	3^{32}	3^8	3^4	1	+1	1	1	5
	$\mathfrak{p}''^2$	$(\bar{\mathfrak{P}}_1\bar{\mathfrak{P}}_2)^2$	1	$_1\;_2\;_3$	1, 1, 1	$\bar{\mathfrak{p}}_1\bar{\mathfrak{p}}_2$	1	2^{36}	1	1	1	+1	1	$\mathfrak{p}^3$	6
	$\mathfrak{p}_1''\mathfrak{p}_2''$	$\bar{\mathfrak{P}}_1\bar{\mathfrak{P}}_2\bar{\mathfrak{P}}_3^2$	1, 1, 1	$_1$ q	1, 1	$\bar{\mathfrak{p}}^2$	1	$2^{12(1+v)}$	$2^{3(1+v)}$	2^{1+v}	2^{1+v}	0	1	1	7

TABLE D

	G	G_Z	G_T	$G_{V_{v_1}}$	$G_{V_{v_2}}$	$G_{V_{v_2+1}}$	$\mathfrak{F}_{\mathfrak{p}}^{(1)}$	$\mathfrak{F}_{\mathfrak{p}}^{(2)}$
1	S	G_8	G_8	G_8	V_4	1	$\mathfrak{p}^{1+v_2}$	$\mathfrak{p}^{1+v_2}$
2	S	G_8	V_4	V_4	1	1	$\mathfrak{p}^{1+v_1}$	$\mathfrak{p}^{1+v_1}$
3	S	V_4	V_4	V_4	V_2	1	$\mathfrak{p}^{1+v_1}$	$\mathfrak{p}^{1+\frac{v_1+v_2}{2}}$
4	S	A	V_4	V_4	1	1	$\mathfrak{p}^{1+v_1}$	$\mathfrak{p}^{1+v_1}$

TABLE E

Case	Possible (v_1, v_2)
$(C, 1)$	(1, -), (3, -), (5, -)
$(C, 2)$	(1, -), (3, -1); (1, -) only when $d \equiv 3$ (9)
$(C, 3)$	(1, -)
$(C, 4)$	(1, -), (2, -)
$(C, 5)$	(1, -)
$(C, 6)$	(1, -), (2, -)
$(C, 7)$	(1, -), (2, -)
$(D, 1)$	(1, 3), (1, 5), (1, 7)
$(D, 2)$	(1, -)
$(D, 3)$	(1, 3)
$(D, 4)$	(1, -)

where $\mathfrak{r}$ is divisible by $\mathfrak{p}^1$, at most. Since $B(\sqrt{\mu}) = B(\sqrt{\bar{\mu}})$, we repeat the argument in the preceding paragraph, since now $m' = h'$.

Every integer in $(B\sqrt{\mu})$ has the form

$$\frac{\rho + \sigma\sqrt{\mu}}{4\mu}, \tag{3}$$

where ρ and σ are integers in B. Hence

$$\frac{\rho + \sigma\sqrt{\mu}}{4\mu} + \frac{\rho - \sigma\sqrt{\mu}}{4\mu} = \frac{\rho}{2\mu}, \qquad \frac{\rho + \sigma\sqrt{\mu}}{4\mu} \cdot \frac{\rho - \sigma\sqrt{\mu}}{4\mu} = \frac{\rho^2 - \mu\sigma^2}{16\mu^2}$$

are integers in B. Hence $\rho = 2\mu\bar{\rho}$, $\sigma = 2\bar{\sigma}$, with $\bar{\rho}$ and $\bar{\sigma}$ integers in B, so that

$$\mu\bar{\rho}^2 \equiv \bar{\sigma}^2 \pmod{4\mu}. \tag{4}$$

The relative discriminant of (3) is $\bar{\sigma}^2\mu^{-1}$. Since the discriminant of $B(\sqrt{\mu}) \mid B$ is the greatest common ideal divisor of all the $\bar{\sigma}^2\mu^{-1}$, we must determine a solution $\bar{\rho}$, $\bar{\sigma}$ of (4) for which $\bar{\sigma}^2\mu^{-1}$ is divisible by a minimum power of $\mathfrak{p}$. We indicate by e the order of $\mathfrak{p}$ in 2 and by V the order of $\mathfrak{p}$ in the discriminant of $B(\sqrt{\mu}) \mid B$, that is, $\mathfrak{p}^e \parallel 2$, $\mathfrak{p}^V \parallel$ discriminant of $B(\sqrt{\mu}) \mid B$.

THEOREM 1. *When $\mu \equiv 0$ (mod $\mathfrak{p}$), $V = 2e + 1$. When μ is prime to $\mathfrak{p}$, V is even, $V = 2V'$, and V' is in fact that integer for which*

$$\mu \equiv x^2 \qquad (\mathfrak{p}^{2(e-V')})$$

has a solution x (not necessarily integral) in B, while

$$\mu \equiv x^2 \qquad (\mathfrak{p}^{2(e-V')+1})$$

is unsolvable.

Let $\mu \equiv 0(\mathfrak{p})$. By (4), $\mu\bar{\rho}^2 \equiv \bar{\sigma}^2 \pmod{\mathfrak{p}^{2e+1}}$. We get, sequentially, $\bar{\sigma} \equiv 0(\mathfrak{p})$, $\bar{\rho} \equiv 0(\mathfrak{p})$, $\bar{\sigma} \equiv 0(\mathfrak{p}^2)$, $\bar{\rho} \equiv 0(\mathfrak{p}^2)$, $\cdots$, $\bar{\sigma} \equiv 0(\mathfrak{p}^{e+1})$, so that $\bar{\sigma}^2\mu^{-1} \equiv 0(\mathfrak{p}^{2e+1})$. Hence $V = 2e + 1$ in this case.

Let μ be prime to $\mathfrak{p}$. Then, since $\bar{\sigma}^2\mu^{-1}$ is divisible by $\mathfrak{p}$ only if $\bar{\sigma}$ is, V is even, $V = 2V'$. Since, by (4), $\mu\bar{\rho}^2 \equiv \bar{\sigma}^2 \pmod{\mathfrak{p}^{2e}}$, $\bar{\rho} \equiv 0(\mathfrak{p}^{V'})$ and[14]

$$\mu \equiv \left(\frac{\bar{\sigma}}{\bar{\rho}}\right)^2 \pmod{\mathfrak{p}^{2(e-V')}}.$$

Conversely, let $\mu \equiv \left(\dfrac{\bar{\sigma}}{\bar{\rho}}\right)^2 \pmod{\mathfrak{p}^{2(e-V')}}$ where $\bar{\sigma}$ and $\bar{\rho}$ are integers in B and V' has the minimum value; since μ is prime to $\mathfrak{p}$, we may assume that $\bar{\rho}\bar{\sigma}$ is prime to $\mathfrak{p}$. If $\mathfrak{p} \parallel \pi$, then $\mu(\bar{\rho}\pi^{V'})^2 \equiv (\bar{\sigma}\pi^{V'})^2 \equiv \sigma_1^2 \pmod{\mathfrak{p}^{2e}}$ and $\sigma_1^2\mu^{-1} \equiv 0(\mathfrak{p}^{2V'})$, so that $V \geqq 2V'$. If $V \geqq 2V' + 1$, then $\sigma_1^2\mu^{-1} \equiv 0(\mathfrak{p}^{2V'+1})$, $\sigma_1 \equiv 0(\mathfrak{p}^{V'+1})$ and $\bar{\sigma} \equiv 0(\mathfrak{p})$, whereas we had taken $\bar{\sigma}$ prime to $\mathfrak{p}$. Hence $V = 2V'$ and the theorem is established.

For what is to follow we must make a digression on the number of incongruent quadratic residues modulo certain powers of $\mathfrak{p}$ that are prime to $\mathfrak{p}$ ("prime" quadratic residues).

THEOREM 2. *Modulo $\mathfrak{p}$ or $\mathfrak{p}^2$ there are $\varphi(\mathfrak{p})$ incongruent prime quadratic residues. If $e \geqq 2$, then modulo $\mathfrak{p}^3$ or $\mathfrak{p}^4$ there are $\varphi(\mathfrak{p}^2)$ incongruent prime residues, whereas modulo p^5 there are $\varphi(p^3)$ such. There exist numbers in B that are prime quadratic residues modulo $\mathfrak{p}^2$ (or $\mathfrak{p}^4$) and are non-residues modulo $\mathfrak{p}^3$ (or $\mathfrak{p}^5$).*

Every integer α prime to $\mathfrak{p}$ gives rise to a prime quadratic residue modulo $\mathfrak{p}$, namely α^2. If α_i ($i = 1, 2, \cdots, \varphi(\mathfrak{p})$) ranges over a complete set of incongruent integers prime to $\mathfrak{p}$, then so do their squares. For, if $\alpha_i^2 \equiv \alpha_j^2 \pmod{\mathfrak{p}}$, then $\alpha_i \equiv \pm\alpha_j \equiv \alpha_j \pmod{\mathfrak{p}}$, since $-1 \equiv +1 \pmod{\mathfrak{p}}$.

---

[14] V' can be taken $\leqq e$, since always there exist solutions of (4) with $\mathfrak{p}^e \parallel \bar{\rho}$, $\bar{\sigma}$.

If, now, β_j ($j = 1, 2, \cdots, N(\mathfrak{p})$) ranges over the $N(\mathfrak{p})$ incongruent residues modulo $\mathfrak{p}^2$ that are $\equiv 0 \pmod{\mathfrak{p}}$, then the $\alpha_i + \beta_j$ ($i = 1, 2, \cdots, \varphi(\mathfrak{p})$; $j = 1, 2, \cdots, N(\mathfrak{p})$) constitute a complete set of $\varphi(\mathfrak{p}^2)$ prime residues modulo $\mathfrak{p}^2$. For a fixed i, all the $\alpha_i + \beta_j$ yield the same quadratic residue modulo $\mathfrak{p}^2$, and $(\alpha_i + \beta_j)^2 \equiv (\alpha_k + \beta_l)^2 \pmod{\mathfrak{p}^2}$ implies $i = k$. For, then $\alpha_i^2 \equiv \alpha_k^2 \pmod{\mathfrak{p}}$, so that again $\alpha_i \equiv \alpha_k \pmod{\mathfrak{p}}$. Hence there are $\varphi(\mathfrak{p}^2)/N(\mathfrak{p}) = \varphi(\mathfrak{p})$ incongruent prime quadratic residues modulo $\mathfrak{p}^2$.

Modulo $\mathfrak{p}^3$ the situation is similar. We now let α_i range over the $\varphi(\mathfrak{p}^2)$ incongruent prime residues modulo $\mathfrak{p}^2$ and let β_j range over the $N(\mathfrak{p})$ incongruent residues modulo $\mathfrak{p}^3$ that are $\equiv 0 \pmod{\mathfrak{p}^2}$. Then, as before, we get $(\alpha_i + \beta_j)^2 \equiv (\alpha_k + \beta_l)^2$ implies $i = k$. For, $\alpha_i^2 \equiv \alpha_k^2(\mathfrak{p}^3)$ and $0 \equiv \alpha_i^2 - \alpha_k^2 \equiv (\alpha_i^2 - 2\alpha_i\alpha_k + \alpha_k^2) + 2\alpha_k(\alpha_i - \alpha_k) \equiv (\alpha_i - \alpha_k)^2 \pmod{\mathfrak{p}^3}$ (since $e \geqq 2$ and $\alpha_i \equiv \alpha_k \pmod{\mathfrak{p}}$ so that $2(\alpha_i - \alpha_k) \equiv 0(\mathfrak{p}^3)$). Hence $\alpha_i \equiv \alpha_k \pmod{\mathfrak{p}^2}$ so that $i = k$. There are, then, $\varphi(\mathfrak{p}^3)/N(\mathfrak{p}) = \varphi(\mathfrak{p}^2)$ incongruent prime quadratic residues modulo $\mathfrak{p}^3$.

In a similar manner, we treat the cases for the moduli $\mathfrak{p}^4$ and $\mathfrak{p}^5$ and the results are those stated in Theorem 2.

It remains to prove the last part of the theorem. If μ is a prime quadratic residue modulo $\mathfrak{p}^2$ and if ξ_i ranges over the $N(\mathfrak{p}^2)$ incongruent numbers modulo $\mathfrak{p}^3$ that are $\equiv 0 \pmod{\mathfrak{p}}$, then the numbers $\mu + \xi_i$ form a set of $N(\mathfrak{p}^2)$ incongruent values modulo $\mathfrak{p}^3$. Since there are only $\varphi(\mathfrak{p}^2)$ incongruent quadratic residues modulo $\mathfrak{p}^3$, there can be at most $\varphi(\mathfrak{p}^2)$ values $\mu + \xi_i$ which are quadratic residues modulo $\mathfrak{p}^3$. Since $\varphi(\mathfrak{p}^2) < N(\mathfrak{p}^2)$, it follows that there exist numbers μ in B such that $\mu \equiv x^2 \pmod{\mathfrak{p}^2}$ is solvable for x in B, whereas $\mu \equiv y^2 \pmod{\mathfrak{p}^3}$ has no solution for y in B.

A similar argument shows that there exist quadratic residues modulo $\mathfrak{p}^4$ that are not quadratic residues modulo $\mathfrak{p}^5$. Theorem 2 is, therefore, completely proved.

By means of Theorems 1 and 2 we can now ascertain which cases in Table C that correspond to 2 as an irregular prime divisor *cannot* exist. We consider the various entries in Table E.

In case $(C, 1)$ according as $v = 1, 3$, or 5, we must have $\mu \equiv x^2(\mathfrak{p}^4)$, $\mu \not\equiv y^2(\mathfrak{p}^5)$; $\mu \equiv x^2(\mathfrak{p}^2)$, $\mu \not\equiv y^2(\mathfrak{p}^3)$; and $\mu \not\equiv x^2(\mathfrak{p})$. As we have already shown, there exists a μ in B for which the first two conditions can be satisfied. $v \neq 5$, for every integer prime to $\mathfrak{p}$ is a quadratic residue modulo $\mathfrak{p}$.

In case $(C, 4)$, $\mathfrak{F}_{\mathfrak{p}}^{(1)} = 1$, $\mathfrak{F}_{\mathfrak{p}}^{(2)} = \mathfrak{p}^{1+v}$ with $v = 1, 2$. The case $v = 1$ is impossible, as every integer prime to $\mathfrak{p}$ is a quadratic residue modulo $\mathfrak{p}$; the case $v = 2$ may, however, be realized.

The discussion of the remaining cases is very similar to that of the first four and need not be gone into further; the application of Theorems 1 and 2 is direct.

4. **Questions of existence of normal fields of degree 24 having the symmetric group.** All normal fields $N \mid R$ having the symmetric group and realizing a given Hilbert subgroup series for a rational prime p are included among the

fields $B(\sqrt{\mu_1}, \sqrt{\mu_2})$ where B is a normal sextic field with the symmetric group of degree 3, in which p has the desired properties, and $B(\sqrt{\mu_1})$, $B(\sqrt{\mu_2})$ are independent quadratic fields over B.

Under what circumstances such a field N is normal[15] with the symmetric group of degree four is a question not answered by the present paper. We can advance somewhat closer to a solution, however, by the following considerations.

If G is the group of such a field N, and if N is normal, then the question immediately arises: *What groups G of order 24 have as subgroups*[16] *S_3 and V_4, with V_4 as an invariant subgroup, and with G/V_4 isomorphic to S_3?*

It can be shown[17] that there are exactly four such groups. These are (a) S_4, (b) $S_3 \times V_4$, (c) $\Phi = \{A, B, Q\}$ where

$$A^4 = B^2 = Q^3 = 1, \qquad BAB = A^{-1}, \qquad A^{-1}QA = Q^{-1}, \qquad BQ = QB,$$

and (d) $= \{A, B, Q\}$ where

$$A^4 = B^2 = Q^3 = 1, \qquad AB = BA, \qquad A^{-1}QA = Q^{-1}, \qquad BQ = QB.$$

The case (b) of the direct product can be eliminated if at least one of μ_1 and μ_2 is not a rational number. But the cases (c) and (d) are not so easily removed from consideration and further study of this question will be necessary to isolate cases (a), (c), and (d), cases corresponding to different types of quartic fields.

PRINCETON UNIVERSITY & THE INSTITUTE FOR ADVANCED STUDY

---

[15] For normalcy alone we need only require that the ideal group H in B corresponding to N as a class field over B be invariant under all the automorphisms of B. But to impose the condition that N shall have, in addition, the symmetric group of degree four as Galois group is a problem of greater depth.

[16] We denote by S_3 and S_4 the symmetric groups of degree 3 and 4, respectively.

[17] The details of this demonstration, while strictly elementary, are not given here, for the proof, employing the known list of groups of order 24, is of a computational nature. These will be given in a later paper of the author.